

University of Redlands

InSPIRe @ Redlands

MS GIS Program Major Individual Projects

Theses, Dissertations, and Honors Projects

8-2009

Using GIS to Support University Development Efforts

Michael Liebler

University of Redlands

Follow this and additional works at: https://inspire.redlands.edu/gis_gradproj



Part of the [Geographic Information Sciences Commons](#), and the [Higher Education Administration Commons](#)

Recommended Citation

Liebler, M. (2009). *Using GIS to Support University Development Efforts* (Master's thesis, University of Redlands). Retrieved from https://inspire.redlands.edu/gis_gradproj/139



This work is licensed under a [Creative Commons Attribution 4.0 License](#).

This material may be protected by copyright law (Title 17 U.S. Code).

This Thesis is brought to you for free and open access by the Theses, Dissertations, and Honors Projects at InSPIRe @ Redlands. It has been accepted for inclusion in MS GIS Program Major Individual Projects by an authorized administrator of InSPIRe @ Redlands. For more information, please contact inspire@redlands.edu.

University of Redlands

Using GIS to Support University Development Efforts

A Major Individual Project submitted in partial satisfaction of the requirements
for the degree of Master of Science in Geographic Information Systems

by

Michael Liebler

Fang Ren, Ph.D., Committee Chair

Diana Sinton, Ph.D.

August 2009

Using GIS to Support University Development Efforts

Copyright © 2009

by

Michael Liebler

The report of Michael Liebler is approved.

A handwritten signature in cursive script, reading "Diana Sinton".

Diana Sinton, Ph.D.

A handwritten signature in cursive script, reading "Fang Ren".

Fang Ren, Ph.D., Committee Chair

September 2009

Acknowledgements

I would like to acknowledge the following people for their contributions and support: Leslie Liebler, David Liebler, Peggy Tuttle, Fang Ren, Sean Pack, Johnson Kosgei, Christopher Loyd, Shane Lehman, and Ray Watts.

Abstract

Using GIS to Support University Development Efforts

by
Michael Liebler

The University of Redlands Development Office seeks to cultivate relationships with alumni to secure charitable gifts to provide the most beneficial learning environment for current and future students. The development office uses several strategies to aid in this process, but lacks an understanding of the geographic component of their donor database. This project will employ an approach that utilizes GIS to reach a more comprehensive understanding of the network of donors the University of Redlands maintains by analyzing spatial statistics within the data. Adding this spatial knowledge component will aid the Development Office in using its resources more effectively and efficiently.

Table of Contents

Chapter 1 – Introduction.....	1
1.1 Client	1
1.2 Problem Statement	2
1.3 Proposed Solution	2
1.3.1 Goals and Objectives.....	2
1.3.2 Scope	3
1.3.3 Methods	3
1.4 Audience.....	4
1.5 Overview: Remainder of the Report	4
Chapter 2 – Background and Literature Review.....	7
2.1 Philanthropy as a Social Science.....	7
2.2 Philanthropy and Higher Education	7
2.3 Geo-Demographics.....	8
2.4 Analyzing Donation Patterns with GIS and Statistics.....	9
2.4.1 Analyzing Geographic Patterns of Donations	9
2.4.2 Modeling Relationships.....	11
2.5 Summary	11
Chapter 3 – Systems Analysis and Design	13
3.1 Problem Statement	13
3.2 Requirements Analysis.....	13
3.3 System Design.....	15
3.4 Project Plan	16
3.5 Summary	17
Chapter 4 – Database Design	19
4.1 Conceptual Data Model.....	19
4.2 Logical Data Model.....	20
4.2.1 Normalization and its Importance	20
4.2.2 Potential Donor Class Relationships	21
4.2.3 Potential Donor, Donation, and Loyalty Class Relationships	22
4.2.4 Potential Donor, Residence, and Demographic Class Relationships	23
4.3 Data Sources.....	24
4.4 Data Scrubbing and Loading.....	26
4.5 Summary	26
Chapter 5 – Implementation	27
5.1 Alumni Data Explorer Development.....	27
5.2 Geocoding and Data Preparation.....	30
5.3 Statistical Methods	32
5.4 Summary	34
Chapter 6 – Results and Analysis	35
6.1 Alumni Data Explorer Functionality.....	35

6.2	Choropleth Mapping	36
6.2.1	Where are Alumni and Donors?.....	37
6.2.2	Does Distance Matter?	44
6.3	Descriptive Statistics	48
6.4	Statistical Results	53
6.4.1	Who are More Likely to Donate?.....	54
6.4.2	Who are More Likely to Donate More?	55
6.4.3	Mapping Targeted Potential Donors in Southern California	58
6.5	Summary	62
Chapter 7	– Conclusions and Future Work.....	63
	Works Cited.....	65
Appendix A.	Mapping Statistics by States	67

Table of Figures

Figure 3.1. System Design	16
Figure 4.1. Conceptual Data Model	19
Figure 4.3. UML Diagram: Composition of Potential Donor	22
Figure 4.4. UML Diagram: Potential Donor, Loyalty, Donation Class	23
Figure 4.5. UML Diagram: Potential Donor, Demographic, Residence Class	23
Figure 5.1. ADO Data Source	28
Figure 5.2. Alumni Data Explorer	28
Figure 5.3. ADE Tool Strip	29
Figure 5.4. Residence Table in Design View	30
Figure 5.5. Geocoding Results	31
Figure 5.6. The Result of a Spatial Join	32
Figure 6.1. ADE and ArcGIS Explorer with Alumni Data	35
Figure 6.2. Residence Table Launched from the ADE	36
Figure 6.3. Total Alumni Map	38
Figure 6.4. Donation Percentages Map	39
Figure 6.5. Total Giving Map	41
Figure 6.6. Average Giving Map	42
Figure 6.7. Top 20 Zip Codes Map	44
Figure 6.8. Total Alumni and Total Lifetime Giving Distance Map	46
Figure 6.9. Percentage that Donate and Avg. Lifetime Giving Distance Map	47
Figure 6.10. All Disciplines Summary Graph	51
Figure 6.11. Bachelors Degree Summary Graph	52
Figure 6.13. Graduate Degree Summary Graph	53
Figure 6.14. ANOVA Mean Plots	58
Figure 6.15. All Alumni Southern California	59
Figure 6.16. Potential Donors Southern California	60
Figure 6.17. Alumni Density Map and Targeted Alumni in S. California	61

List of Tables

Table 3.1 System Requirements for Prototype Data Explorer.....	14
Table 3.2 System Requirements for Alumni Database.....	15
Table 4.1 Data Characteristics.....	25
Table 6.1 Total Alumni Top Ten.....	37
Table 6.2 Percentage of Donating Alumni.....	39
Table 6.3 Total Lifetime Giving Top Ten.....	40
Table 6.4 Average Lifetime Giving Top Ten	42
Table 6.5 Top 20 Zip Codes	43
Table 6.6 Percentiles Summary of Donating Alumni Lifetime Giving.....	48
Table 6.7 Greek Membership Summary.....	49
Table 6.8 Years Since Graduation Summary.....	50
Table 6.9 Summary Table for all Alumni Disciplines.....	51
Table 6.10 Summary Table for Alumni with a Bachelor's Degree.....	52
Table 6.11 Summary Table for Alumni with a Graduate Degree.....	53
Table 6.12 Explanatory Variables.....	54
Table 6.13 Independent Samples T-Test Results.....	55
Table 6.14 OLS Results.....	56
Table 6.15 One-Way ANOVA Results.....	57
Table 6.16 Example of Four Targeted Alumni.....	62

List of Acronyms and Definitions

ADE	Alumni Data Explorer
AGX	ArcGIS Explorer
ADO.net	Active Data Object
ANOVA	Analysis of Variance
GUI	Graphic User Interface
OLEDB	Object Linking and Embedding Database
OLS	Ordinary Least Squares
VB	Visual Basic
VMR	Variance Mean Ratio

Greek Household Abbreviations

ACD	Alpha Chi Delta
AGN	Alpha Gamma Nu
APA	Alpha Phi Alpha
ASP	Alpha Sigma Pi
ATP	Alpha Theta Phi
AXO	Alpha Xi Omicron
BLM	Beta Lambda Mu
CRP	Chi Rho Psi
CSC	Chi Sigma Chi
DKP	Delta Kappa Psi
KSS	Kappa Sigma Sigma
PBS	Phi Beta Sigma
PIC	Pi Chi
PST	Phi Sigma Tau
SKA	Sigma Kappa Alpha
ZPB	Zeta Phi Beta

Chapter 1 – Introduction

Institutions of higher learning depend on donations from foundations, industry, the community, and especially alumni to provide financially for current and future students. Monetary gifts are supremely important to the survival of a university. Thus, many approaches have been devised and studied regarding the topic of philanthropic giving, the motivations behind gift giving, and how to maximize an organization's ability to receive these gifts. Currently, the University of Redlands maintains complete records of its donors and its Development Office intends to gain more insight into the gift giving behavior of their alumni by analyzing these records.

To accomplish this goal, this project sought to provide spatial awareness and analysis to the previously tabular donor records. Specifically, a geocoded donor database was developed and a set of spatial statistical analyses of gift giving behavior was conducted. In doing so, this project demonstrated how a geographic information system (GIS) can serve as an efficient tool to help the University of Redland's Development Office maximize its limited resources by cultivating relationships with existing and future alumni.

Chapter one of this project is organized into five sections. The client section provides a brief history of The University of Redlands, as well as an introduction to the client for this project. Following an introduction to the client is a synopsis of the problem the University Development Office faces, as well as a proposed solution to this problem. Specifically, the goals of the project and the scope that these goals fall into are discussed in the solution section. The overall scope of the project is then followed by a methodology subsection that seeks to provide an outline of the procedures completed to achieve the project's goals. Chapter one concludes by clarifying the intended audience of this project and a summary of the organization of the remaining report.

1.1 Client

The University of Redlands, a private college located in San Bernardino County California, was founded in 1907 by members of the American Baptist Church. The first class at the university consisted of 59 students. The university has maintained a strong relationship with the Baptist Church. Mandatory attendance to church service was only eliminated in 1972 (Redlands, 1999).

The College of Arts and Sciences was the first formal college established at the university, and remains a cornerstone of the university's academics. An individualized study college known as the Johnson Center was added in 1979. This unique approach to learning allows students to design their own bachelor's degree program. The university continues to grow its College of Arts and Sciences opportunities and has added additional educational opportunities for a more diverse learning environment, including the College of Education and College of Business.

The prospect of enhancing the university through fundraising was initially established in 1928. The first person to be given the title of "Alumni Director" was Arthur Jacobsen. Responsibilities of the Alumni Director included alumni relations, the giving program, and alumni activities and events. The first alumni event was held in

1950. By 1954, the number of staff set for fund raising efforts included one director, and three support employees. At present the University of Redlands Development Office consists of a director and nearly 20 support staff (Chandra 2007).

The individual client within the University of Redlands for this project is Ray Watts, the director of the University of Redland Development office. Mr. Watts has introduced several staff members within the development office to aid in the project, and provided insight into the existing donor data. He has served as a knowledge base for questions about the data as well as the overall goals for the project.

1.2 Problem Statement

The University of Redlands Development Office maintains a rudimentary donor database in a tabular format. The problem with the tabular database is that geographic distribution of donors and potential relationships between donors and their geographic context cannot be revealed easily. Without a solid understanding of the spatial component of gift giving behavior, the development office cannot maximize its gift receiving potential across space. Therefore, it was determined that the implementation of a solution that used spatial analytical methods in a GIS would help solve this problem.

1.3 Proposed Solution

Providing one all-encompassing solution to the issue of enhancing the University Development Office's donor program was not a realistic endeavor. The following section discusses the overall goals of the project and the methods used, paying special attention to the GIS created for the University of Redlands development office and how it will offer a more complete understanding of the donor database. Beyond the raw numbers within the donor database, the spatial information derived via spatial statistics, and spatial analysis will provide greater insight into the donor profile that should be potentially targeted by the staff of the University Development Office.

1.3.1 Goals and Objectives

While simply displaying the donor database on a map would be somewhat helpful, the ultimate goal is to accompany the raw data provided by the University of Redlands Development Office with a complete array of relevant statistical and spatial analyses. Under this goal, there are several objectives. First, the raw donor data should be displayed on a map. With the donors being visualized on a map, the data becomes more conducive to using GIS technology for further analysis. The second objective of the project was to explore the existing donation spatial patterns, and provide a comprehensive profile of the donors that not only is determined by donor's attributes, but also may be relevant to the geographic contexts.

The final objective of the project was to design and propose GIS solutions to allow the office staff to enter, update, visualize, and explore the donor data. This will greatly aid the decision making process in place within the University Development Office. Ultimately, a new mechanism for the University of Redlands will be introduced to achieve more success in their efforts to raise capital to provide for future students.

1.3.2 Scope

Solving the problem of spatial awareness within the University of Redlands Development Office began by geocoding the existing raw donor data. The geocoding process required the establishment of a level of precision at which to geocode the individual alumni records. For the purpose of this project it was determined that the zip code level would be an appropriate level of accuracy because of the large volume of records. The zip code attribute was also nearly completely free of entry errors, while there are many missing values for the finer geographic scale such as street numbers. Correspondingly, the statistical analysis on the data was done at a zip code scale.

In order to create a comprehensive donor profile, the relevant geographic indicators for the characteristics of the residential areas needed to be identified. As such, obtaining additional geographic data was important for a successful analysis of the alumni network. The most updated demographic data from 2008 was provided by Environmental Systems Research Institute (ESRI), from which additional variables were determined and used in the statistical analysis conducted on the alumni network.

In addition to displaying the alumni on the map and linking them with the relevant variables from the demographic data, a major task was to examine the gift giving behavior of Redlands' alumni using spatial statistical methods. Before beginning with more complex statistical methods it was important to provide a comprehensive descriptive statistical analysis of the alumni within the donor network. Descriptive statistics such as percentiles, averages, and totals were an important exploratory measure that helped dissect the data prior to conducting more advanced statistical techniques. The scope of the statistical analysis only reflects to the basic hypothesis testing methods given the time frame of this project.

While the spatial statistical analysis was an important component of this project, implementing a longer term GIS solution would be ideal for the University of Redlands Development Office. The first component of the GIS solution was a prototype Microsoft Access database. This database contains completely normalized data in which each attribute pertaining to alumni at the university is placed in the proper table and related to each other table via an alumni identification number. The second component was a standalone interface that allowed the user to interact with the data within the database without opening the database and compromising the data. A prototype user interface was included in the project scope.

Finally, the limitations of the analyses completed on the report were identified, and additional development opportunities are discussed. The opportunities for further project development included a discussion on a complete GIS implementation plan, further development of the database GIS application, and additional statistical analysis methods.

1.3.3 Methods

This section will discuss the multi-step process used to prepare the alumni data for analysis, including geo-enabling the donor data, compiling statistics of the alumni network, and completing the prototype database, and interface used to illustrate a long-term GIS solution.

When received, the donor data was organized into an Excel spreadsheet in a rudimentary fashion. To display the tabular data on a map, the first method used in the project was geocoding, which is the process that converts address texts into geographic coordinates. Before geocoding, the data went through several processing iterations. All attendees of the University of Redlands who did not graduate were eliminated from the spreadsheet. Records containing an address outside the United States (297 alumni records) were also removed. A default address locator provided by ESRI was then utilized to geocode the remaining alumni. This process successfully matched 99 percent of the alumni within the spreadsheet. A shapefile was then created containing all attributes determined to be appropriate for statistical analysis.

After the donor data was geocoded, the University of Redlands Alumni shapefile was then spatially joined to state polygons and zip code polygons. The spatial join technique allowed additional summation attributes to be added to the newly-created polygon's table of attributes. The resulting polygons were of identical shape as the original input; however, they now contained useful descriptive statistics derived from the original individual donor data.

Charts, graphs, and relevant maps were then created to illustrate items such as totals, averages, minimums, maximums, and standard deviations of various attributes, providing a synopsis of the overall distribution of the data on a nationwide scale. Following the exploratory phase, the increasingly advanced analysis performed utilized several statistical techniques to define variables that affected donations. The analysis also sought to study any previously indiscernible patterns that were present in the data. In the end the analysis provided a more complete understanding of which alumni donate, the variables that affect these donations, and the variables that affect donation amounts.

1.4 Audience

This report is intended for use by non-GIS professionals. The report was requested by the University of Redlands Development Office Director, who has some basic knowledge of basic GIS techniques; however, the Director has not yet incorporated any spatial database or GIS software into the daily work flow. The audience is familiar with the University of Redlands, the University's needs, and the University's long term goals. It may also shed light on the similar studies for other institutions.

1.5 Overview: Remainder of the Report

The remainder of this report is organized into six chapters. Chapter Two provides a review of articles and books relevant to the goals of this report. Chapter Three details the design of the project to be implemented in the report. This chapter includes discussions on the evolution of the system design, and the resources that were dedicated to making this a successful project.

Chapter Four delves further into the system design specifically that of the database suggested to the University Development Office. The chapter provides details about the data, their sources, and the preparation of the data. Chapter Five discusses the procedures to successfully implement the proposed GIS solutions identified in Chapters Three and Four. It also expands on the methodology section introduced in Chapter One.

Chapter Six provides the results of the statistical and spatial analyses of the alumni data. The chapter presents several maps and tables intended to visually summarize the data in a format useful to the client. The final chapter provides a summary of the client's requirements, how these requirements were satisfied, and identifies several ways in which the project could be expanded upon in the future.

Chapter 2 – Background and Literature Review

Literature on the subject of philanthropy is abundant, and many have taken a systematic approach to understanding what motivates people to give. Literature discussing proven strategies on securing donors to fund non-profits provides a psychological approach but few specifically tackles the subject of using GIS and or spatial statistics to support the procurement of charitable gifts. More often researchers have used methods such as linear regression to find explanatory variables to account for or help dissect a particular issue.

2.1 Philanthropy as a Social Science

As described in *Strategic Giving: The Art and Science of Philanthropy* (Frumkin 2006), understanding philanthropy beyond the fundamental question of how much a person has donated is essential to grasping the multiple aspects of the subject. Payton and Moody (2008) study real world examples of philanthropic giving, and use theoretical explanations to provide an understanding for the rationale behind a donation. Philanthropy can be studied methodically, and viewed as a social science.

Over time, cultures develop a unique philanthropic tradition that reflects the society in which they live. The United States, for example, is a conglomeration of ethics and morals from throughout the world. Payton's discussion of "the social history of moral imagination" (p. 131-133) is an informative discussion first introduced by anthropologist Clifford Geertz. Simply speaking, moral imagination is used to describe and determine what is important in a particular culture or society. John Payton expands on the concept of moral imagination by stating that to understand philanthropic traditions, one must know that "philanthropy is the social history of the moral imagination because philanthropic action in any time or place requires the use of moral imagination by the philanthropist" (Payton, 2008, p. 133).

The process of securing gifts for the stability of any non-profit organization has shifted over time and become more strategic and business-like in its methods. Payton discusses terms such *venture philanthropy*, *strategic philanthropy*, and *new philanthropy*. New philanthropy seeks to provide for causes in which a donation will provide actual success rather than just a quick fix, a much more business-like philosophy. Although philanthropic giving has become increasingly like big business, it's important to note that the majority of money secured by non-profits is still from the individual (Payton, 2008). This fact underlies the importance of analyzing available data such as an alumni network on an individual basis.

2.2 Philanthropy and Higher Education

Charitable giving to both public and private universities is paramount for the ongoing success of higher learning institutions. As emphasized by Charles Vest (2006), "Despite such interesting and important challenges, the generosity of individuals and foundations and the support of far-sighted corporations and industry consortia are central to maintaining and enhancing America's outstanding system of higher education." Vest's article dissects the importance of gift giving to support financial aid, scholarships,

research, and the overall excellence of universities. Vest (2006) further emphasizes the importance of gift giving and endowment income in the success of private colleges and universities. One notable fact he mentions is 22 percent of all private American Association of University Schools income comes from private gifts and grants. Only 20 percent of revenue is generated by tuition (Vest, 2006).

The article proceeds by continuing to discuss why the success of universities, particularly private universities, is so dependent on individual donors, predominantly extremely wealthy donors. Vest claims that maintaining the excellence of a university depends on the priorities of the wealthiest Americans. This claim relates back to the concept of social imagination mentioned earlier in the review.

One of the challenges discussed in the article is the problem universities face as they seek to balance the needs of the university with the vision the donor had when providing the gift. Issues including bureaucracy and state and federal regulations, among other issues all play a role in the large gift giving process, as well (Vest, 2006). Again the importance of a universities donor network is apparent. The quality of education for current and future college and university students relies heavily on gift giving and corporate sponsorships.

In addition to recognizing the motivations, models, theoretical explanations, and the social scientific aspect of charitable giving, it's important to understand the role a geographic information system can play in enhancing a complete understanding of the subject.

2.3 Geo-Demographics

The donor database provided by the Development Office at the University of Redlands contained records of both alumni and non-alumni that have given to the university. ESRI (2002) describes the importance of expanding upon these records to create a more sophisticated database that enables the user to better understand donor behavior. Combining demographic data with an accurate geographically enabled database can aid in identifying the area of greatest asset accumulation, and provide further insight into the relationship between charitable gifts and where the donor lives.

Expanding upon geo-demographics, Harris (2005) explains that the ability to use different classification techniques on US census data will allow for a more complete household profile. Harris goes on to discuss how geo-demographic data will provide the capacity for in-depth research on neighborhoods and a further understanding of how to classify neighborhoods in which existing donors to the University of Redlands live.

Grubestic (1997) discusses how to generate geo-demographics from an organization's internal data. His methods and strategies may be used on the data provided by the University of Redlands. Grubestic further expands upon the creation of low cost geo-demographics, and their sources. By incorporating this additional data within the University of Redlands donor database, Grubestic suggests researchers will be able to identify lifestyle cluster groups. Each relevant piece of additional data added to the existing donor database will potentially provide a much more comprehensive analysis in both descriptive and variable statistical studies of the University of Redlands donor network.

Geo-demographics are also often used with GIS technique to reveal how the subject of interest, such as donation, relates to the different demographic attributes across

space, such as median household income of the neighborhoods. This group of studies will be further discussed in 2.4.1.

2.4 Analyzing Donation Patterns with GIS and Statistics

The use of GIS to study patterns and statistics is a principle component of this project. ESRI (2002) describes GIS as a useful tool to not only geographically enable a donor database, but also to study the spatial statistics of the donor database. Utilizing GIS can help an organization, such as a university, to identify spatial patterns of the donation and/or the relationships within the donor database that were previously indiscernible. Therefore, the utilization of GIS can further support the University of Redlands decision making processes when soliciting and securing donations.

2.4.1 Analyzing Geographic Patterns of Donations

By utilizing the power of human visions, mapping has been demonstrated to be a useful way for spatial data exploration, hypothesis generation, and knowledge construction (MacEachren et al. 2004). Particularly, with the advancement in GIS techniques, mapping has become much easier and various maps can be created using available GIS software. To date, research has been conducted using GIS to examine the geographic patterns of donation in different contexts.

One example completed by the Office of Institutional Research at Binghamton University used GIS specifically to analyze alumni distribution and donation patterns to enhance the targeting and marketing efforts of the University (Jardine, 2003). In this study, an alumni dataset composed of 75,973 records was first geocoded at zip code level with a 98% success rate. The geocoding procedure assigned geographic coordinates to each alumnus, which made mapping the alumni feasible.

Several choropleth maps were created to depict alumni totals, total alumni giving, average alumni giving, and percentage at which alumni donate within each state. Each of these statistics was also thematically mapped at the county level within the State of New York. These efforts showed that 60% of alumni resided in New York, and aided the institution in refining donor targeting efforts. Furthermore, the study attempted to examine the relationship between average donation amounts and median household income by overlaying the graduated symbol map of average giving onto choropleth maps of the median household income. The mapping result does not reveal the clear relationship between median household income and the amount of donation at county level for the State of New York.

The final mapping output sought to illustrate the potential voting influence of current Binghamton University students and alumni by showing their totals within congressional districts in the State of New York. This demonstrated how GIS could potentially be used to influence policy and federal funding. The methods utilized within this particular article showed how descriptive statistics and thematic mapping could be used to gain a greater understanding of alumni locations and donation patterns.

Remote Control is another example of utilizing mapping to illustrate donor contributions (MAPlight.org, 2008). This study used choropleth mapping to illustrate the percentage of funds a US house member received outside of their congressional district

from January 2005 to December 2007. Of the 421 house members studied, 97 percent received over half of their funds outside of the actual district they represent. Graduated symbol maps for individual house members created for the study showed the distribution of their funding. The report concludes with summary tables showing the top 100 zip codes to have donated to any house member, followed by a breakdown of individual house members and the contributions they have received in and out of district. This study provided insight into the relationship between distance and donation patterns by showing that a close proximity does not necessarily increase contributions to an entity.

Another study by Gimpel (2006) investigates how geographic location and social status may influence a potential donor's campaign contributions to the Democratic and Republican Party. The author explains how *spatial proximity* plays a significant role in the decision to contribute. A close proximity facilitates communication, and facilitates organizing around a common interest.

Using data from the U.S. Federal Election commission, Gimpel further examined individual contributions to the two major political parties. This data was analyzed on a zip code scale by aggregating individual contributions to their zip code of origin. The spatial distribution of the contributions was then mapped with GIS software for both political parties, which provided insight into the distribution of campaign funds received across the United States. Using demographic variables such as population density, the study identified the degree to which each party was dependent on contributions from metropolitan areas. Median household income was utilized to show the dependence each party has on contribution funds received from wealthier neighborhoods. Each party was shown to have a heavy dependence on these contributions although these wealthy zip codes only represented 18% of the total voting population.

In summary, mapping and summarizing distributions of donors can be accomplished in various fashions which all involve geoenabling the data being studied. Once this data is geoenabled, GIS can be a useful tool to create summary tables and various mapping products to illustrate information with thematic mapping. Overlaying project data with additional demographic data is a useful method when seeking to gain insight on donation behavior.

Besides various mapping techniques, there is another group of methods for spatial pattern analysis which focuses on testing the null hypotheses that a spatial pattern has been generated by complete spatial randomness process. These methods are useful for finding significant geographic patterns. Several procedures can be employed to test the degree of randomness of a spatial pattern. Rogerson (2006) discusses methods for spatial point analysis, which include quadrat, nearest neighbor analysis, and Ripley's K function. These methods can aid in identifying a clustering or a dispersed point pattern. In the aspect of areal data, spatial autocorrelation presents as the clustering or dispersion of attributes of areal units. Spatial autocorrelation is an important issue to taken into account because it violates the independence assumption for many standard statistical techniques, including T-Tests, ANOVA test, and regression analysis. Several methods can be exercised to test the degree of spatial autocorrelation within a dataset at both global and local levels (Getis and Ord, 1992; Anselin, 1995; O'Sullivan & Unwin 2002). Moran's I is the traditional global measure of spatial autocorrelation (Moran 1950) while LISA is the local measure for detecting local spatial clusters and outliers (Anselin, 1995).

2.4.2 Modeling Relationships

Following the spatial analysis of sophisticated donor database, there are several different statistical analysis methods that can be utilized to provide a further understanding of the donor database. One fundamental method is regression analysis to identify the relationships between various variables. The linear regression models include bivariate and multivariate linear models. Bivariate regression analysis seeks to analyze the relationship between an event (dependent variable) and an explanatory variable (independent variable). For example, the price of a house can be result of the square footage of the house. The square footage, however, is not the only factor in determining housing price. Therefore, a multiple regression technique would be utilized to identify additional independent variables responsible for the variability in housing prices, such as number of bathrooms and age. The regression technique has been extensively used in many studies. For example, Chandra (2007) used three regression models to analyze the relationship between three dependent variables which included the lifetime giving amount, giving over a specified number of years, and an alumni loyalty score. These variables were analyzed against several independent variables. The independent variables used in his study included age, Greek membership, and degree awarded. Chandra concluded that age at the time of graduation and Greek membership were the key explanatory variables in his study. His conclusions were based on using only a small sample of the data.

Unlike the regression on non-spatial variables, the regression on geographic data needs more consideration. Mitchell (2005) discusses how to choose the proper attributes (variables) within GIS datasets to analyze, and the methods that can be used to discover relationships, patterns, and clusters. It's important to account for special considerations when analyzing spatial statistics. The most common problems include the modifiable areal unit problem, boundary problems, sampling issues, and recognizing spatial autocorrelation as mentioned in 2.4.1.

To address spatial autocorrelation issue, there are two basic models that can be used: spatial lag and spatial error models. For example, Gimpel (2006) employed a spatial regression technique to analyze contribution of American voters to the Republican and Democratic Parties. The model applied included a *spatial lag* coefficient within the regression model. This *spatial lag* was calculated by using the eight nearest zip codes within a 60 mile radius as a weight within the model. This technique was used to include the influence of network effects on contributions. From this analysis, Gimpel concluded that there is a positive correlation between prosperous neighborhoods with high resources and contribution amounts. He also suggested that the Republican Party is less dependent on local contributions than the Democratic Party. The study concludes with the findings that the spatial context of voters significantly influences their contribution amounts.

2.5 Summary

Creating a GIS of the existing donor database, and providing spatial analysis of this data will serve as a useful tool for the future goals of the University. Gaining basic knowledge on the theories behind philanthropy and how philanthropy is an integral part of a university's well being illustrates the importance of a study on the University of

Redlands donor network. A comprehensive study can be completed by incorporating relevant demographic data and utilizing methods such as linear regression.

Chapter 3 – Systems Analysis and Design

A detailed system analysis and design were conducted in the early phase of the project to address the problem discussed in Chapter One. This chapter further confirms the problem present within the University of Redlands Development Office, provides in detail the requirements of the client via a table, and discusses the system design. Finally, this chapter outlines the multi-phase process of implementing a longer-term GIS solution.

3.1 Problem Statement

The client faced two issues when understanding the donor network of the University. First, the client does not utilize any structured database models for maintaining their donor network. The donor network data was originally provided by the client in a Microsoft Excel Spreadsheet, which did not allow the development office staff to easily organize, query, analyze, and maintain its entire donor network. Second, the lack of geographic context within the donor data does not allow the spatial component of donor network to be revealed. Gift receiving potential was not able to be examined across space. Therefore, the development office would benefit from an implementation of a relational database that can easily translate into GIS analysis.

3.2 Requirements Analysis

A requirements analysis was done to determine the needs of University of Redlands Development Office. The first need identified was a tool that would allow a user to manipulate data within the alumni database. Specifically, the functional requirements of the tool are to allow a user to update, delete, query, or explore the database (Table 3.1). The functional requirements of the tool could be satisfied using a form created directly in Microsoft Access; however, the user would have to open the database to reach this form, which might not be ideal. Therefore, a standalone interface, the Alumni Data Explorer (ADE), was designed to allow the user to access the alumni database without opening the actual database. In doing so, this explorer was able to provide ease of use, custom functionality, and to preserve the integrity of database.

Table 3.1 Prototype System Requirements for Data Explorer

Functional Requirements of Alumni Data Explorer		
Item	Description	Remark
Update Data	Updating data is a functional requirement of the Alumni Data Explorer Tool. Its purpose is to allow the user to change existing data or enter new data into the University Alumni Database.	Mandatory
Delete Data	Deleting data utilizing the Alumni Data Explorer is a functional requirement of the tool. A user can delete existing records, or delete erroneous data within a record.	Mandatory
Query Data	Querying the data is a functional requirement of the Alumni Data Explorer. This tool will be used to quickly retrieve a record.	Mandatory
Explore Data	Exploring the data refers to the functional requirement of navigating through individual records.	Mandatory
Launch GIS Client	Custom buttons located on the toolbar allow for the user to easily launch a mapping client such as AGX.	Mandatory

The second need identified was the development of a spatial database to serve underlying the Alumni Data Explorer. Broadly speaking, the spatial database seeks to accommodate both spatial and non-spatial data in a manner that allows the user to more effectively organize alumni records. Table 3.2 provides a brief discussion on the three functional requirements of the database that would serve the University of Redlands Development Office. These functional requirements include organizing the alumni data, facilitating data manipulation, and preserving the integrity of the data.

Table 3.2 Prototype System Requirements for Alumni Database

Functional Requirements of Alumni Database		
Item	Description	Remarks
System for Monitoring and Organization	The functional need of the database is to provide an efficient method for organizing and manipulating donor data.	Mandatory
Overall Database Requirements	Preserve Data Integrity.	Mandatory
Performance	The searchable fields within the database are indexed to improve query and sorting speeds.	Desired

Another goal stipulated by the client was to gain insight into how donors' gift giving varies across space, for the purposes of effectively targeting future potential donors in the nation. Specifically, the functional requirement for this component of the project includes enhancing the client's understanding of the spatial distribution of donors and helping the client to identify the relationships between geographic indicators and gift giving behaviors. Although this component of the project is not part of the Alumni Data Explore application, it complements the data exploration by providing a useful spatial analysis.

3.3 System Design

Based on the functional requirements discussed above, the project system was designed to manage data via data entry, data navigation, data editing, and data display and visualization. The system design for the University of Redlands Development Office was straight- forward. Since there was no database or GIS component currently being utilized within the office, this was an initial attempt to provide the client with a standalone graphical user interface (GUI) to update the database. Figure 3.1 provides an illustration of the system design, which includes data input, database development, and a GIS application that includes custom tools. The design also illustrates the analysis portion of the project that includes mapping outputs and various statistical analyses on the donor data.

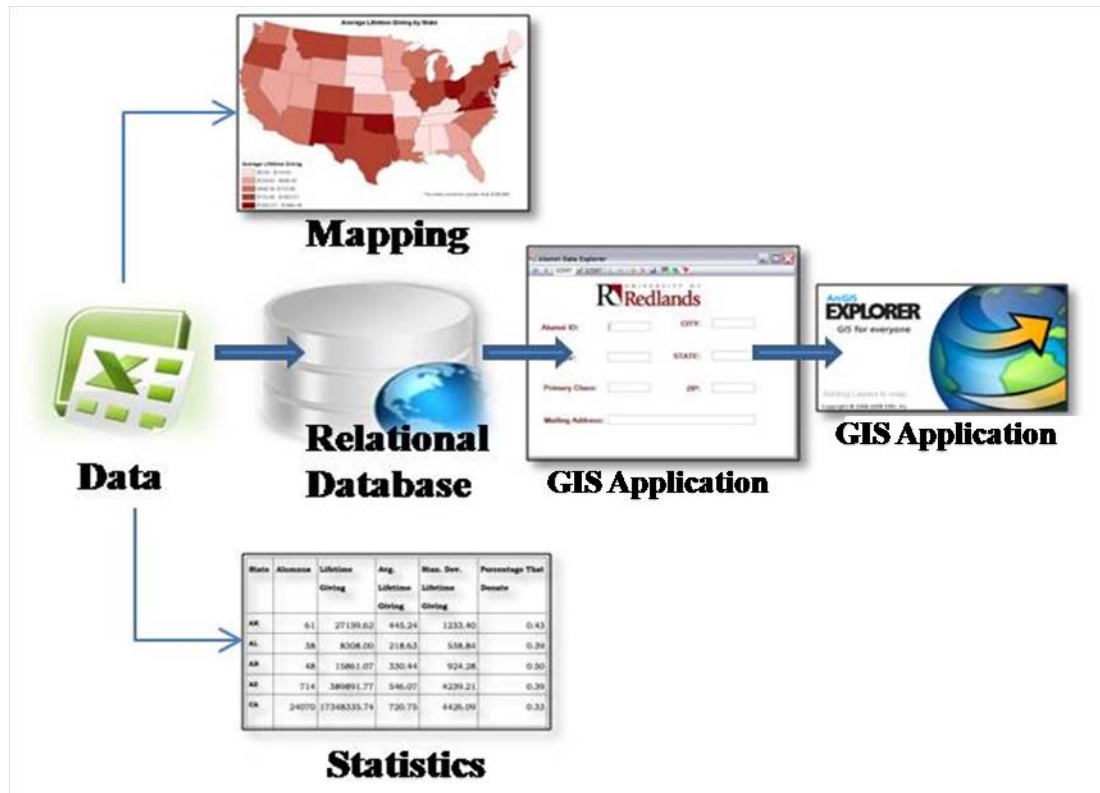


Figure 3.1. System Design

3.4 Project Plan

In the early stages of this project, the goal was simply to understand what the University Development Office wanted as an end product. After developing this understanding the next goal was to understand what the client actually needed and how GIS could help in this endeavor. When all was said and done, it was concluded that the client could not only use what he asked for, (a spatial statistical analysis looking for patterns within his donor data), but also a more long term GIS solution. This solution would include five distinct phases:

Phase I: Conceptualize and design a prototype relational database system

The goal of phase I was to take a convoluted spreadsheet containing records for over 38,000 University of Redlands Alumni, and develop a blueprint and methodology for reorganizing and incorporating this data into a relational database model. This was completed successfully through careful planning using UML diagrams. The relationships, tables, and data characteristics established here led directly to phase II.

Phase II: Create a prototype relational database

Phase II involved ingesting the actual donor Excel Spreadsheet into Microsoft Access. Once the data was ingested, the relationships and appropriate tables were

created. Redundant data was eliminated and every attempt to normalize the database was made. Included in phase II was the plan to identify data within ESRI's business analyst dataset that could be utilized later in the statistical analysis portion of the project.

The most important aspect of phase II was the process of geocoding the donor addresses to an accuracy level appropriate to both analyze and display the data. It was determined that because of the large number of records in the database, and due to the scale of the analysis being conducted, using an address locator that utilized only zip codes would be sufficient. After establishing this threshold, the data was successfully geocoded. This provided a spatial context in the form of x and y coordinates within the database.

Phase III: Create a prototype spatial data explorer

Phase III of the project was created to provide an even more complete GIS solution to the client, in addition to the database and analysis phases. In the early conceptual phases, the Alumni Data Explorer was meant to serve as a standalone tool that would allow a user to manipulate data within the Microsoft Access database without actually opening the database. In addition, the Explorer was to provide a tool to export donor records directly into a mapping client such as ESRI's ArcGIS Explorer (AGX).

Following research into the functionality provided within free mapping clients such as AGX, it was determined that a direct data export function would be beyond the scope of this project. The completed prototype data explorer would include most of the original functionality, including the stand alone interface to navigate the database. To replace the original export donor records' functionality, a custom button was provided to simply launch AGX. The alumni data would be pre-prepared and loaded into ArcGIS Explorer for use by the University Development Office.

Phase IV: Conduct a spatial statistical analysis

When the initial approach was laid out, phase III consisted of various statistical analyses. The first method was used to gain a greater understanding of the data on a nationwide scale. A descriptive statistical analysis was conducted to establish this understanding. A series of maps and charts were created to illustrate the distribution of alumni donations within each state.

The second method involved a more complex statistical analysis using a two sample independent T-Test within SPSS software, independent proportions test, a one-way ANOVA test, and an ordinary least squares analysis. These final four statistical methods allowed for the further exploration of the most telling variables within the alumni data, and the incorporation of three additional demographic variables from ESRI's 2008 demographic dataset.

3.5 Summary

Determining the requirements and needs of the University Of Redlands Development Office proved to be a subjective task. While the initial goal was simply to study relevant statistical approaches to identify areas the sales staff of the development office should target, it was determined that the current approaches utilized by the office could be greatly enhanced by the implementation of a relational database and a tool to explore this

database within a mapping environment. Employing simple GIS components, such as a relational database, and a mapping client such as ArcGIS Explorer would add a missing spatial component to the development office's resource allocation and decision making process.

Chapter 4 – Database Design

Chapter Four provides information on the conceptualization, design, and implementation of the University of Redlands donor database. It first describes the conceptual database model and the components of the database design for this project. Following the conceptual database model is the logical database model, which identifies the relationships between the objects contained in the conceptual model. Finally, this chapter looks at the data characteristics and methodology used to prepare and load the data into the relational database.

4.1 Conceptual Data Model

A conceptual data model is the first-step abstraction of the real world when planning a database. This conceptualization phase identifies the entities within the data and the relationships between each entity, which can be depicted by creating a conceptual schema. The conceptual data model also helps identify and contain the scope of the database. Based on an assessment completed on the data within the University of Redlands donor records spreadsheet, the following conceptual model was created. The conceptual model illustrated in Figure 4.1 represents the entity donor as the main entity which all other entities are related back to.

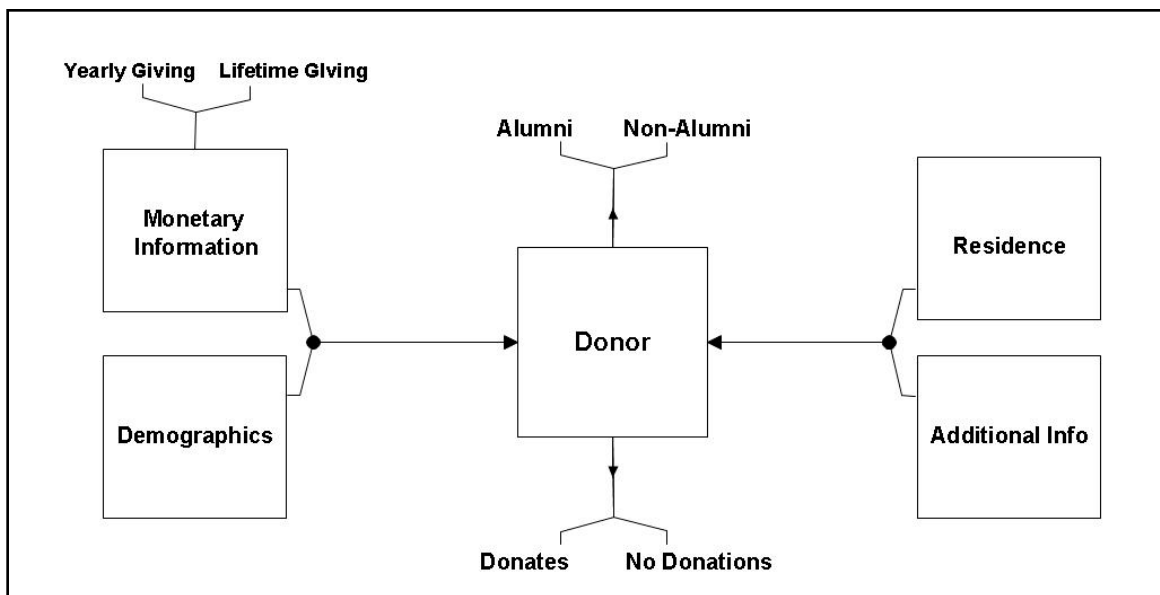


Figure 4.1. Conceptual Data Model

The donor entity can be identified as a person who has either graduated from the university or has not graduated from the university. The donor entity can then be identified as a person who has, or has not donated to the university. Finally, each donor has additional information that provides a clearer picture of each donor record, including

donation amounts, demographic data, location, and other information on a donor's relationship to the University of Redlands.

4.2 Logical Data Model

The relational data model is a popular choice when storing data in a table format because it is conducive to storing spatial and non-spatial data. For this reason, the prototype database design for the University of Redlands Development Office was created using a relational database approach. Every effort was made to normalize this alumni network data.

4.2.1 Normalization and its Importance

One of the keys to the successful long term management of data within a database is to ensure that the database contains normalized data. There are several levels of normalization, each increasing the overall fluidity of the database. When a database is properly designed, it is said to be normalized, or in normal form. Normalization is a process that ensures that data redundancy has been minimized or eliminated, and that data structure remains consistent and properly organized. When normalizing a database, there are three rules that each predicate on each other. The rules are:

- 1st Normal Form
- 2nd Normal Form
- 3rd Normal Form

The first rule of normalization involves ensuring each attribute field contains its smallest meaningful value. For example, an attribute titled *City* would contain only an entry for the name of a city, not the city name and the state that the city is in. The field titled *State* would be an additional field. Fields that are created should be atomic, along with the values keyed in.

The second normal form pertains to tables that contain a multiple-field key. To ensure a database is in second normal form, all non-key fields within individual tables that have multiple-field keys must be pertinent to the multiple-field key. If information contained in one of the multiple-key fields indicates the value of a non-key field within the table, the non-key field should be moved to another table. This eliminates redundancy within the data. If values within a non-key field do not relate back to the multi-key primary field, that non-key field should also be moved to another table. When these two characteristics are satisfied, the database is said to be in second normal form.

The third normal form is achieved by ensuring that all non-key fields within a table relate properly to the key-field. If a field does not relate to the key-field, it should be moved to another table. For example, a table titled *Donation* would not be in third normal form if it contained the field *Zip*. The field *zip* is not an important relationship to the donation table, and should be removed or relocated to the proper table. Ensuring a database meets the rules of normalization is paramount in achieving a well designed database. Following the normalized UML diagram (Figure 4.2 below), the figure has been divided into three sections identifying separate, but overall related class associations.

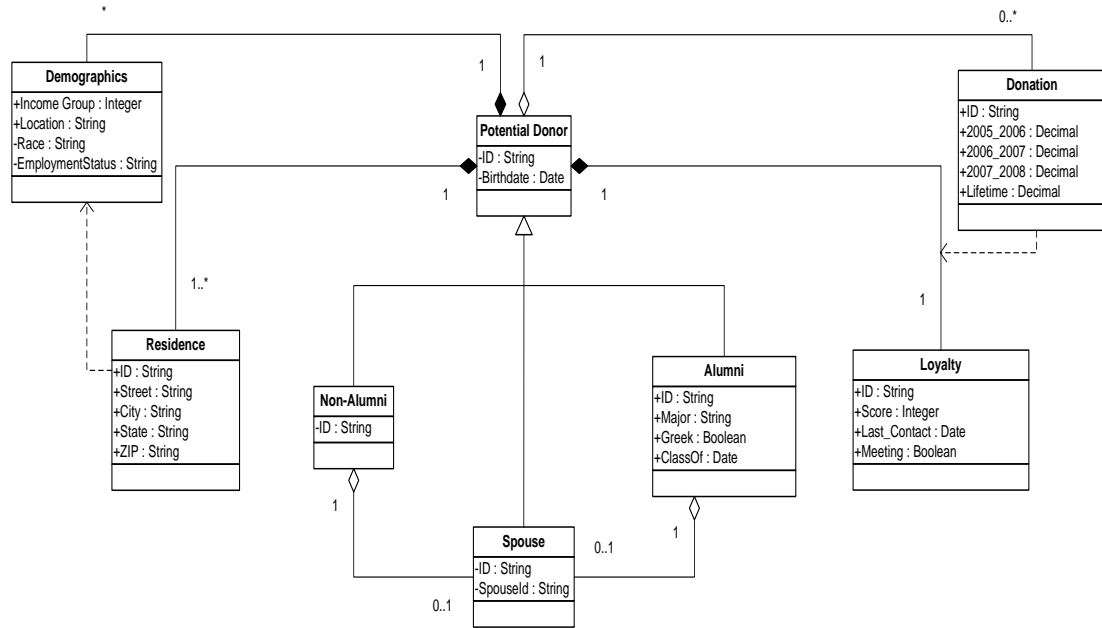


Figure 4.2. Complete UML Design

4.2.2 Potential Donor Class Relationships

The potential donor class is a super class composed of alumnus, non-alumnus, and spouse. Every potential donor has an entry for identification number and date of birth. The alumnus and non-alumnus attribute may have or may not have a spouse who could be a potential donor. Therefore, a loose association exists between spouse and alumnus as well as between spouse and non-alumnus. The attributes for alumnus are ID, major, Greek (Fraternal) association, graduation date, and degree awarded. The ID indicates the alumnus record and will be used as the primary key to relate the class to other classes.

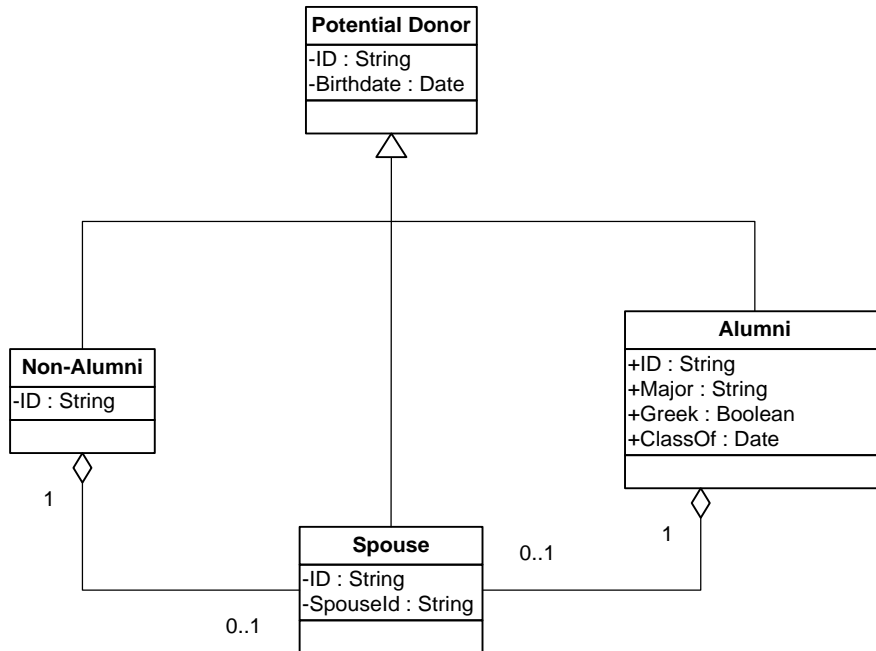


Figure 4.3. UML Diagram: Composition of Potential Donor

The Greek attribute indicates whether the alumnus was involved in a fraternity or sorority. This could give an indication of the degree to which the alumnus feels dedicated to the university, a hypothesis to be explored in Chapters Five and Six. The year of graduation indicates how long ago a particular alumnus graduated. For non-alumnus, the attributes are the identification number and the year of affiliation to the university. The affiliation year refers to any years that the potential donor attended the University of Redlands.

4.2.3 Potential Donor, Donation, and Loyalty Class Relationships

All potential donors are rated by a loyalty score derived over time. This score is based on donation patterns and how involved the potential donor is with University affairs and events. A loyalty score is also a rating of the potential willingness to support the university over a long period of time. The frequency and amount of a donor's contributions increases the loyalty score of a given donor. This score is used to focus advocacy and fund raising resources in an effort to attract more money from these donors. Willingness to donate is manifested through participation of the potential donors

in both non-fundraising, as well as fundraising events.

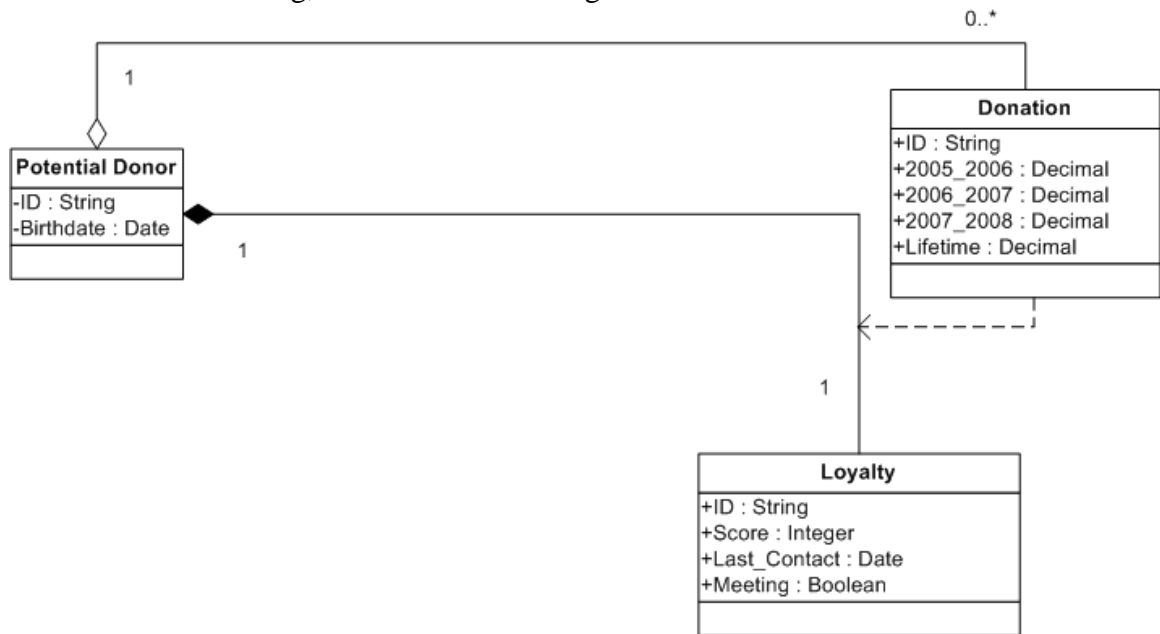


Figure 4.4. UML Diagram: Potential Donor, Loyalty, Donation Class

4.2.4 Potential Donor, Residence, and Demographic Class Relationships

The university also maintains information about the residency of potential donors, as modeled by the residence class. The primary purpose of this class is to provide the university with necessary contact information on potential donors, such as city and zip code. This class and its associated attributes are important to the implementation of this project, since they provide the necessary geographic locations for mapping and analysis

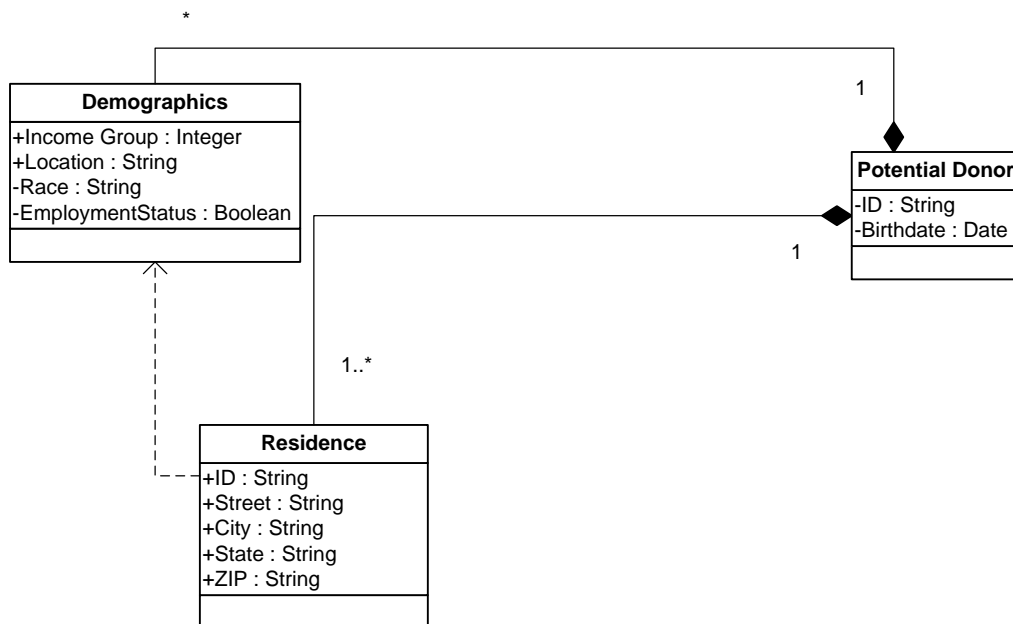


Figure 4.5. UML Diagram: Potential Donor, Demographic, Residence Class

Associated with the residence class is the demographic class. The university does not currently maintain potential donor information specifically for this class. However, some of the demographic class attributes can be harvested from other sources maintained at the university (e.g., age from birth date, employment status, and location from address). As with the residence class, the demographic class is an essential component for the implementation of this project because the University's Development Office is keenly interested in how demographics affect donations. Demographics are strongly spatial in nature, providing more impetus for applying geographic information systems tools and techniques for the purpose of increasing monetary support for the University of Redlands.

4.3 Data Sources

In the early stages of the project, data was provided by the client in a single Microsoft Excel Spreadsheet. The table contained over 38,000 records of University of Redlands alumni, their spouses, and individuals who had completed a semester or more at the university but did not graduate. Each donor record had 18 associated attributes.

An MIDL or master input data list is a summary document that provides details about the data. This summary includes:

- Name and ID number
- Source
- Volume
- Format
- Cost
- Identifies additional work the data will require

A complete MIDL is beneficial when setting up a database. There are four major components of an MIDL. The components are:

- Data identification details
- Data volume considerations
- Data characteristics
- Source data availability & cost

The following MIDL contained in Table 4.1 contains information specific to the setup of the alumni database for the University of Redlands Development Office (Tomlinson, 2007).

Table 4.1. Data Characteristics

COMPONENT	Details Needed	Notes
1. Data Identification	Potential Donor Database	This is the only source of data that will be included in the final project; see item 3 for additional details.
2. Data Volume	Although there are approximately 38,000 people presently in the database, the size in terms of storage space is relatively small.	Data volume & storage is not a problem.
3. Data Characteristics	<p>There is one data source; however the single source can be broken down thoroughly. The data contains several major components that form relationships with each other.</p> <ul style="list-style-type: none">• Alumni ID & Address• Greek Affiliation, Alumni Event Attendance• Graduation Year• Yearly giving,• Lifetime Giving• Mean Giving• Loyalty Score	A complete UML diagram of the relationships within the data has been created for reference.
4. Data Availability & Cost	The data contained within the database was provided free of cost by the University Development Office for education purposes.	Cost & availability are not an issue in regards to this report.

Another set of 2008 demographic data for the U.S. was provided by ESRI and several variables were extracted to represent the geographic contexts of the residential areas of the alumni. The values extracted from the data included the percentage of population who are Caucasian, the percentage of population who have bachelor degree's or higher, and the median household income. These variables were measured at the zip code level.

4.4 Data Scrubbing and Loading

The loading and preparation of alumni data for the report was a principle component of this project. Using the logical data model as a blueprint, the single Excel table was compiled into multiple tables, then ingested into the relational database software Microsoft Access.

Preparing the donor data for geocoding, map making, and analysis involved more steps which are discussed further in Chapter Five. Following the creation of a single shapefile containing all records using a geocoding process, multiple shapefiles were created to facilitate different forms of map making and analysis.

4.5 Summary

The creation of a useful relational database for the University of Redlands Development Office began with the conceptualization of the entities present within the data, and their relationships. Following the conceptualization phase, a relational database was chosen as the data model that would best serve the data provided by the development office. A database schema was created via a UML diagram depicting each table, along with each table's attributes to be created within the relational database software. Subsequent to the UML schema, tables were created within the relational database software, and the potential donor records were integrated from a spreadsheet into the proper tables in the database.

Chapter 5 – Implementation

Chapter Five discusses the methodology undertaken to achieve the final results for this project, and the processes and steps needed to implement similar efforts in future work. The development of the Alumni Data Explorer will first be discussed, followed by the process of geocoding the donor network data. Secondly, the methodology and usefulness of spatially joining data will then be discussed, which will lead into a discussion of the mapping methods and outputs used to illustrate the University of Redlands Alumni Network. Chapter Five concludes with a discussion of the final statistical methods utilized within the project.

5.1 Alumni Data Explorer Development

Creating a stand-alone application to access a relational database is an effective method for exploring the database and preserving the database integrity for non-technical users. A GIS application, called the Alumni Data Explorer (ADE), was developed for the client. While this is only a first phase prototype, the University of Redlands Development Office could certainly benefit from the simplicity of ADE data entry application. The first attempt to develop the application used the VB.net programming language. VB.net proved to be a cumbersome programming method for this situation because the application contained a multitude of similar subroutines in order to incorporate the data entry and edit functionality. To alleviate this problem, an alternative method of programming was chosen.

Following the initial attempt using VB.net, the prototype ADE was created using an Active Data Object (ADO.net) form template within Microsoft Visual Studio 2008. This was an ideal way to create custom applications specifically for managing database systems because it is known as the “standard model for database programming within Visual Studio” (Halvorson, 2008, pg. 438).

The process of creating the Alumni Explorer began by establishing an Object Linking and Embedding Database (OLEDB) connection to the alumni database. This was accomplished by using the data configuration wizard within Visual Studio. After a database connection was created, a blank windows form was added to the Visual Studio

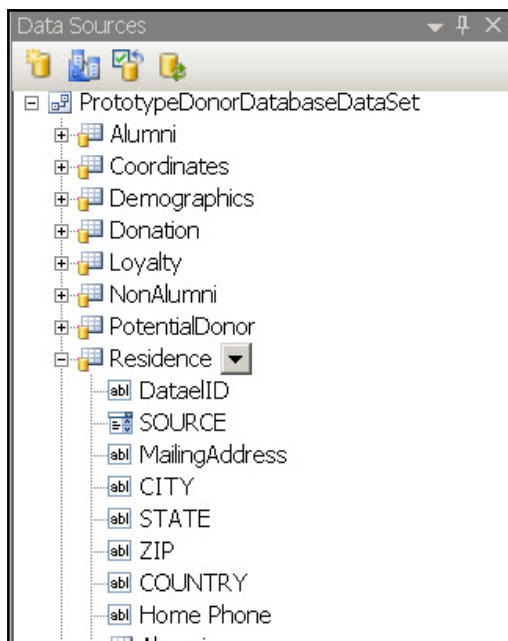


Figure 5.1. ADO Data Source

Project. ADO.net then enabled the dataset designer, a tool which allows easy creation of bounded objects on the newly created windows form in which an interface was designed. The interface was designed to include attributes contained within the residence table. By opening the data source explorer within Visual Studio (Figure 5.1), ADO.net enabled the functionality to drag and drop data sources from tables located in the data source explorer directly onto the windows form; this created entry text boxes. The final prototype application shown in Figure 5.2 illustrates the completed ADE interface that allows for data entry into the residence table within the Microsoft Access Database.

Figure 5.2. Alumni Data Explorer

There are several additional functions contained within the ADE. Located at the top of the form is the button tool strip (Figure 5.3). Beginning from the left, this tool strip contains multiple buttons used to manipulate and navigate through data. This portion of the tool strip required no additional programming, as it was preprogrammed onto the form template upon creation. The complete functionality of the tool strip is expanded upon in Chapter Six.



Figure 5.3. ADE Tool Strip

The two buttons located on the far right of the ADE tool strip were created by custom programming. The second button from the right uses a *shell* command to launch the mapping client, ArcGIS explorer (AGX), by referencing its executable path name on the local disk. When AGX is launched, the client then has the ability to navigate and explore alumni data (zip code level only), as well as to click on individual points to obtain attribute information about a particular alumnus.

The final button, located on the far right of the tool strip, allows the client to launch the full residence table contained in the alumni database. This functionality was accomplished using a *data grid view* on a second Windows Form in Visual Studio, and utilizing the residence table as the loading data source. Once the table has been launched, the client can sort and edit data within the form containing the residence table data. Figure 5.4 illustrates the empty residence table in the design view within Visual Studio 2008. By setting the residence table as the data source for the *data grid view*, the attribute headings are automatically populated on the form.

	DataelID	SOURCE	MailingAddress	CITY	STATE	ZIP
*						

Figure 5.4. Residence Table in Design View

5.2 Geocoding and Data Preparation

Geocoding was the first step for a spatial analysis of the alumni network as it provided the geographic location coordinates of the alumni. The process of geocoding the alumni records began with choosing a scale at which to geocode the data. Because of the large number of alumni records, the data was geocoded at the zip code level. The second task was to choose an address locator; in an effort to avoid unwanted errors, an address locator built in to the ESRI software titled *Zip Code* was chosen to perform the geocoding process.

Once the original Excel Spreadsheet containing the alumni records was loaded into the address locator, the process then created X and Y coordinates for each record. These X and Y coordinates allowed for the creation of a shapefile that could then be loaded into GIS software, such as ArcMap, and further analyzed. The geocoding results were 99% successful following the elimination of data entry errors from the original donor spreadsheet. One such error involved problems with locations on the east coast. These eastern shore zip codes contained a zero as the first digit causing the address locator to reject the zip code by categorizing these records as a mismatch, and excluding these records from the geocoded output. These particular alumni records required manual manipulation within Microsoft Excel before they could be transferred into the address locator. Figure 5.5 illustrates the geocoding results of the alumni records.

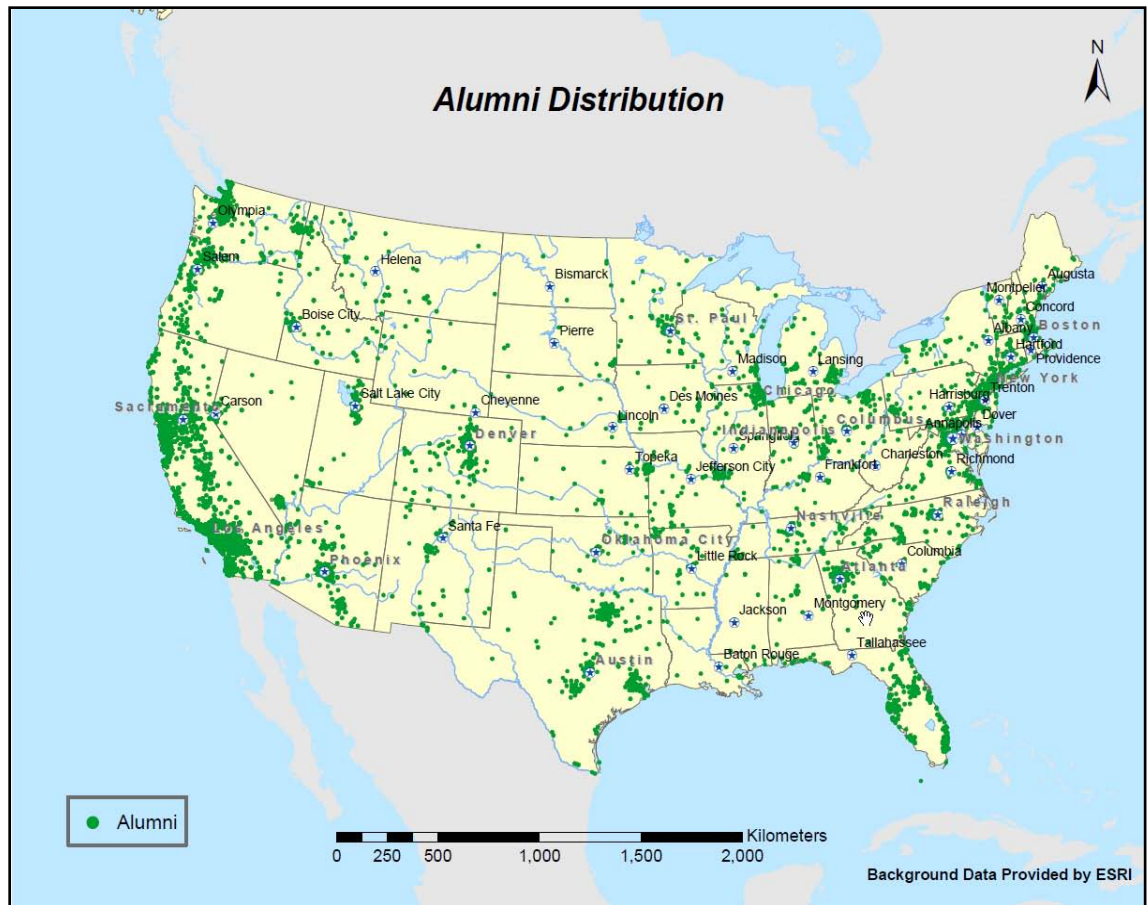


Figure 5.5. Geocoding Results

Several additional data preparation procedures were used to scrutinize the data and prepare it for further analysis. First, the alumni data outside of the United States were eliminated. Second, alumni records with a lifetime giving of greater than \$100,000 were removed from the alumni shapefile for the statistical and spatial analysis. The rationale for removing these donors is that the University Development Office is aware of the presence of these alumni, and has a strong relationship with these individuals. Furthermore, by removing these extreme outliers, the data distribution was less skewed by a few wealthy donors. The final phase of data preparation involved coding variables within the data. Two variables, Greek membership and donating alumni, were converted to and calculated as binary variables, with 1 representing the alumni that participated in a Greek donated and donated to the University while 0 representing the opposite situation.

With geocoding and data preparation complete, the next phase in the project involved using a technique within ArcMap known as spatial join. This function allows point data, such as the alumni records, to be joined to a polygon in which they are located. The polygon choice can be states, counties, or zip codes. When performing the spatial join function within ArcMap, the software also calculates some useful statistics for the point feature aggregation (Figure 5.6). For example, the total count of the points in each polygon and the sum of the designated attribute of the point feature will be

computed. Finally these newly generated attributes are combined with the original polygon attributes to form a new polygon shapefile. Based on this new polygon feature, thematic mapping techniques can be applied to further examine the distribution of the alumni, amount of donations, and other useful attributes.

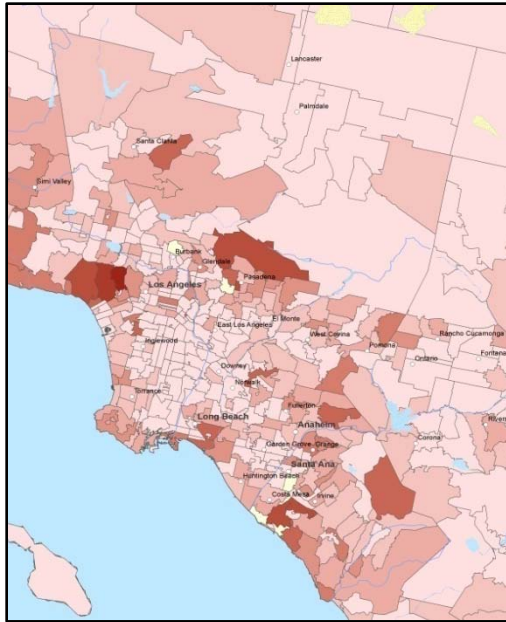


Figure 5.6. The Result of a Spatial Join

Using the completed spatial joins function, useful choropleth maps illustrating the various descriptive statistics generated with the spatial join function were then created. The mapping outputs marked the first time that the University of Redlands Development Office was able to view the alumni network on a map. These maps displayed various statistics such as total alumni, total lifetime giving, average lifetime giving, and the percentage of alumni that donate in a particular region (e.g. state, county, or zip code).

5.3 Statistical Methods

ArcMap and Excel were utilized to generate summary tables for various regional and individual variable statistics. Using the *summarize* tool within ArcMap, tables illustrating the total, mean, , and percentage of donating alumni within a region were created. During the completion of these descriptive statistic tables, particularly close attention was paid to the variables pertaining to Greek membership and graduation time. Since these variables had been classified into binary data, summarizing totals and percentages was simple. Chapter Six discusses the results of these tables.

The next phase of the statistical analysis required the existing donor data to be appended with additional data. This supplementary data was provided by ESRI's 2008 demographic data, which contained hundreds of socio-economic variables that aid in business decision making by creating a donor profile with the neighborhood statistics. To further expand the understanding of both who donates, and how much is donated, three

variables pertaining to the neighborhood characteristics were added to the alumni shapefile: percentage of individuals who have a bachelor's or higher degree (PercEdu); percentage of the population who are Caucasian (PercWhite); and median household income (MedHHInc). Each variable represented a neighborhood statistic at the zip code level, and was used for the statistical tests to come. Four statistical methods were used for the final analysis portion of the project:

- Independent samples T-Test
- Two sample proportions test
- Ordinary least squares test
- One-Way ANOVA

The independent samples T-Test was the first test to explore the concept of how alumni who donated differed from those who did not. Five characteristics of a donor profile were considered, including two individual variables and three neighborhood variables. The individual variables were Greek membership and graduation time. The neighborhood variables were the three mentioned above (PercEdu, PercWhite, and MedHHInc). This procedure was completed with SPSS and was used to test whether these two groups of alumni significantly differed in these perspectives. Since the Greek membership variable is a binary variable, the two samples proportions test was also applied. By comparing the proportion of Greek members within the donate group to the proportion of Greek members within the non-donate group, the hypothesis that Greek members are more likely to donate than non-Greek members was tested.

To answer what factors affect the amount of donation, the ordinary least squared (OLS) procedure was used to test if the five variables are relevant to average lifetime giving. It was expected that alumni who graduated earlier, participated in Greek membership, and live in a relatively affluent, highly educated, and white-dominated neighborhood are more likely donate more. ArcMap was employed to create an output table which included the level of significance for each variable, and the overall model fitness of the OLS test (r-squared value).

However, due to the severely skewed data, the OLS model was not solid, and therefore the one-way ANOVA test was employed to compare whether alumni who fall in different categories of donating alumni are significantly different from each other. Completed with SPSS, the procedure used a factor variable (average giving) and compared the variance of means of four additional variables (PercEdu, PercWhite, graduation time, medHHInc) identified as dependent on this factor variable. To complete this test, the factor variable was first grouped into five categories and assigned a numerical value based on average donation as follows:

- \$0-\$50 =1
- \$50-\$100=2
- \$100-\$200=3
- \$200-\$500=4
- \$500 and above=5

With the average donation amounts categorized, the one-way ANOVA test created table outputs that illustrated the mean of each variable for each categorical average donation value. ANOVA also provided a significance value for each dependent variable.

5.4 Summary

The implementation of the Alumni Data Explorer was accomplished by using programming techniques specifically designed for database management. With ADO.net, a first phase application was created to manipulate data within the alumni database. Geocoding the data required some data preparation, including elimination of entry errors, and manual manipulation of entries beginning with a zero. Preparing the data for analysis was a multi-step process involving eliminating outliers, quantifying nominal data into binary values, and aggregating point data into polygons via the spatial join technique. Following the creation of choropleth maps, various descriptive data summary tables were generated. The final implementation of the project involved using supplemental socio-economic neighborhood statistics to further analyze the characteristics on the donating and non-donating alumni.

Chapter 6 – Results and Analysis

6.1 Alumni Data Explorer Functionality

The completed Alumni Data Explorer (ADE) prototype was a success in that the application performs all the functions it was meant to demonstrate, while not opening the actual database. Utilizing the preprogrammed buttons on the tool strip, the application can successfully navigate, enter, update, and delete records within the Microsoft Access Database. In addition, custom buttons were successfully programmed on the top tool strip. One such button launches ArcGIS Explorer (AGX).

While working in AGX, the client can now navigate throughout a mapping environment with the alumni point data overlaid with various maps, or high resolution imagery. The alumni point data is populated with the various attributes found within the shapefiles that were used for analysis. Figure 6.1 shows alumni point data and a single point's associated attributes.

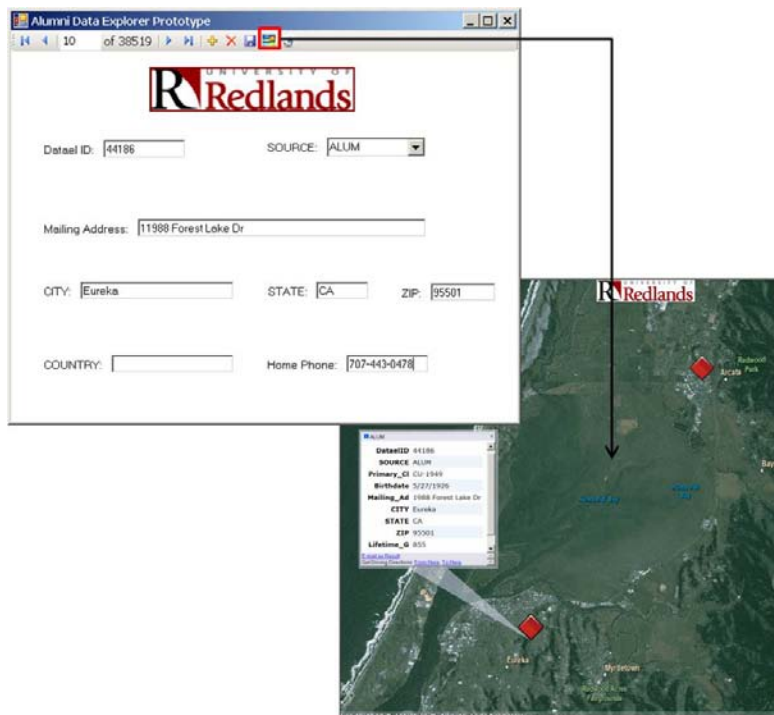


Figure 6.1. ADE and ArcGIS Explorer with Alumni Data

As mentioned in Section 5.1, there was an additional custom tool developed on the far right of tool strip that launches the residence table from the actual database. With this button, the client can sort records by various attributes, edit records, or confirm entry of a new record. Figure 6.2 shows the populated residence table launched from the Alumni Data Explorer.

	DataeID	SOURCE	MailingAddress	CITY	STATE	ZIP
▶	265	ALUM	PO Box 20291	Stanford	CA	94309-02
	266	ALUM	913 W Fern ...	Redlands	CA	92373-58
	274	ALND	808 Balboa Ct	Redlands	CA	92373
	275	ALND	6756 Tiara ...	Highland	CA	92346-60
	277	ALUM	1231 N Gilb...	Fullerton	CA	92833
	278	ALUM	26455 N Ro...	Valencia	CA	91355
	280	ALUM	420 Burlingt...	Long Beach	CA	90803
	281	ALUM	8462 Mesa ...	Santee	CA	92071
	283	ALND	13559 Frem...	Yucaipa	CA	92399-55
	284	ALUM	1009 Meado...	West Covina	CA	91792-10
	285	ALUM	33370 Rose...	Yucaipa	CA	92399-64
	309	ALUM	921 Sunset ...	Redlands	CA	92373-69
	313	ALUM	9756 Diamo...	Yucaipa	CA	92399-29
	315	ALUM	3821 Runni...	Bakersfield	CA	93311-22
	354	ALUM	11005 Quee...	Bakersfield	CA	93312-70
	424	ALUM	10 Osprey	Irvine	CA	92604-45
	435	ALUM	1863 Midnig...	Beaumont	CA	92223-84
	437	ALUM	1526 Chalu...	Davis	CA	95618-67
	438	ALUM	326 San Raf...	Redlands	CA	92373-49
	456	ALUM	304 W Las ...	Altadena	CA	91001-43
	457	ALUM	4430 Banne...	Long Beach	CA	90807
	460	ALUM	264 E Fern ...	San Bern...	CA	92404

Figure 6.2. Residence Table Launched from the ADE

The ADE serves its purpose as an illustration of the usefulness of a standalone data manipulation application; however, the functionality of the tool is limited as it is programmed only to access the data of one table in its present state. There are no queries programmed into the application, and any direct navigation would only be accomplished by knowing the record number of an alumni and typing the number into the navigation text box. In order to develop this application further, additional programming and set-up of relationships between tables is necessary. With the abundance of tables within the database, the creation of more than one application may be necessary.

6.2 Choropleth Mapping

Choropleth mapping is a useful means to represent data on a map by shading data via classification of a particular variable based on its respective values. Capitalizing on the spatial join technique discussed in Chapter Five, the following section uses this mapping methodology on various statistics generated from the alumni dataset. In addition to each map, a supplementary table holding relevant data pertaining to each map has been included.

6.2.1 Where are Alumni and Donors?

The alumni data was spatially joined to states, and several choropleth maps were created to visualize the distribution of alumni and their donations. Figure 6.3 presents the spatial distribution of total number of alumni throughout the entire country. Table 6.1 lists the top ten states represented on this map. The complete table can be found in Appendix A. The total alumni map illustrates the total alumni within each state by using five classification groupings. The west coast has the most alumni, with California containing over 65% of total alumni within the data.

Table 6.1. Total Alumni Top 10

State	Alumni	Lifetime Giving (\$)	Avg. Lifetime Giving (\$)	Percentage That Donate (%)
CA	24155	17,348,335.74	720.75	33
WA	859	552,455.47	643.14	45
AZ	714	389,891.77	546.07	39
OR	636	485,978.61	764.12	44
TX	563	51,1375.32	908.30	37
CO	544	407,033.99	748.22	41
NV	356	124,847.42	350.70	35
HI	326	203,366.22	623.82	53
VA	277	452,033.02	1631.89	44
FL	242	102,564.84	423.82	31

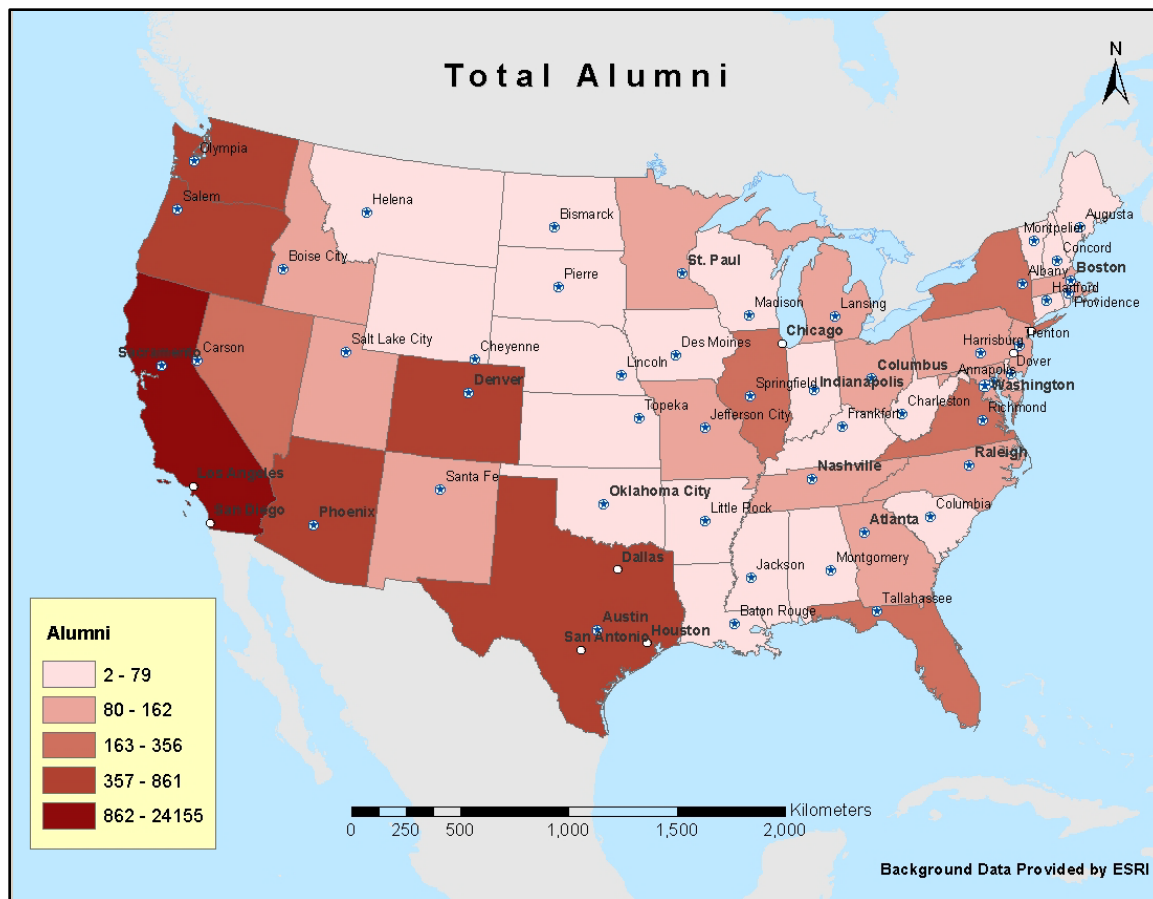


Figure 6.3. Total Alumni Map

Although the west coast has the most total alumni, the percent at which this region donates is near the bottom. The majority of the high donation percentage states are in the east. The issue with percentage is that it is very sensitive to the total number of alumni in each state. For example, the high donation percentage of North Dakota is the result of low total alumni (seven records). Table 6.2 presents the states with the highest donation percentage, while Figure 6.4 illustrates the percentages at which alumni donate.

Table 6.2. Percentage of Donating Alumni Top 10

State	Percentage That Donate (%)	Alumni	Lifetime Giving (\$)	Avg. Lifetime Giving (\$)
ND	71	7	4,009.93	572.85
NH	65	23	10,402.20	452.27
DE	63	8	3,792.00	474.00
HI	53	326	203,366.22	623.82
DC	53	49	83,188.65	1697.73
MA	52	133	167,283.15	1257.77
IA	52	44	21,969.52	499.31
NJ	51	91	123,431.21	1356.39
WI	50	68	3,9473.39	580.49
AR	50	48	1,5861.07	330.44

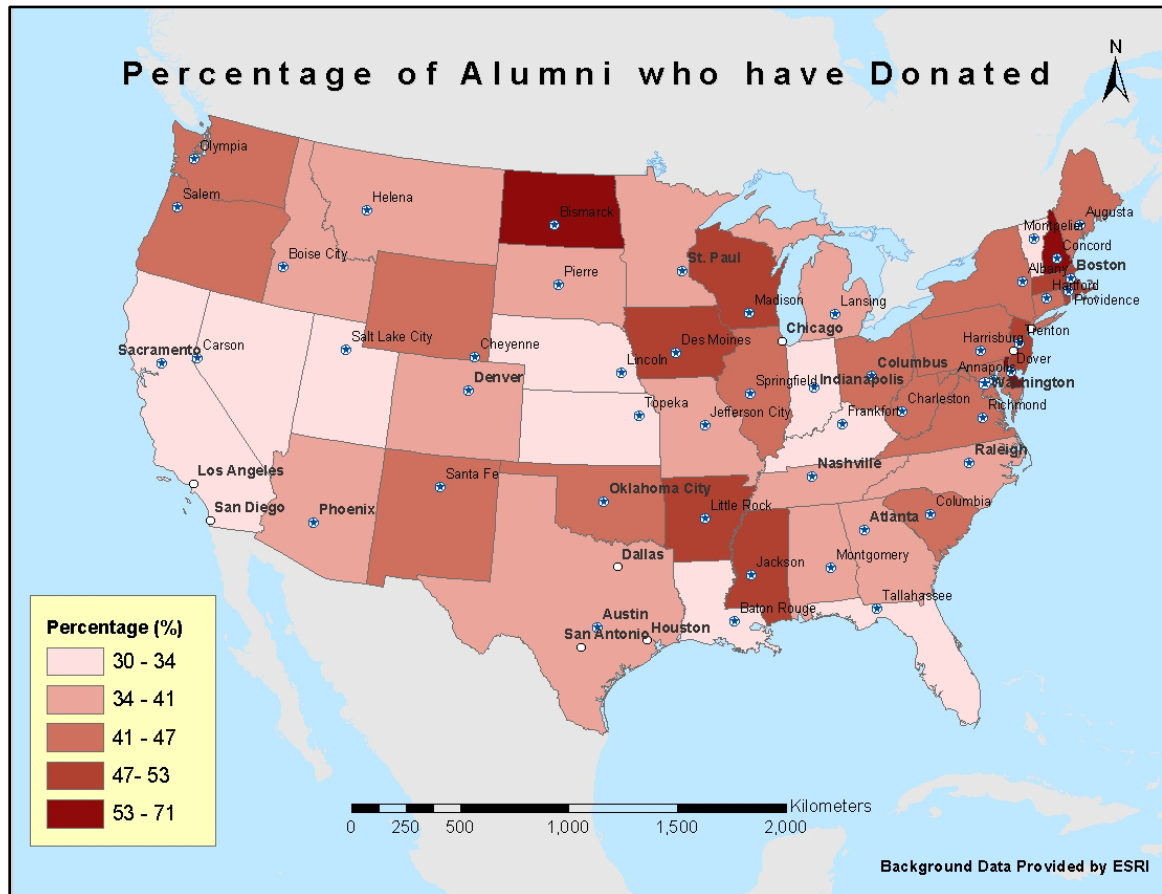


Figure 6.4. Donation Percentages Map

As expected, with the majority of alumni being in and around California, the total lifetime giving is greatest in this region (Figure 6.5). Virginia is a notable state, ranking fifth in total lifetime giving, despite its distance from Redlands, CA. The relationship between distance and donating is further examined in Section 6.2.2. Table 6.3 lists the top donating states in order of lifetime donation amounts.

Table 6.3. Total Lifetime Giving Top 10

State	Lifetime Giving (\$)	Alumni	Avg. Lifetime Giving (\$)	Percentage That Donate (%)
CA	17,348,335.74	24,070	720.75	33
WA	552,455.47	859	643.14	45
TX	511,375.32	563	908.30	37
OR	485,978.61	636	764.12	44
VA	452,033.02	277	1,631.89	44
CO	407,033.99	544	748.22	41
AZ	389,891.77	714	546.07	39
NM	273,161.99	150	1,821.08	47
OH	228,217.64	124	1,840.46	43
HI	203,366.22	326	623.82	53

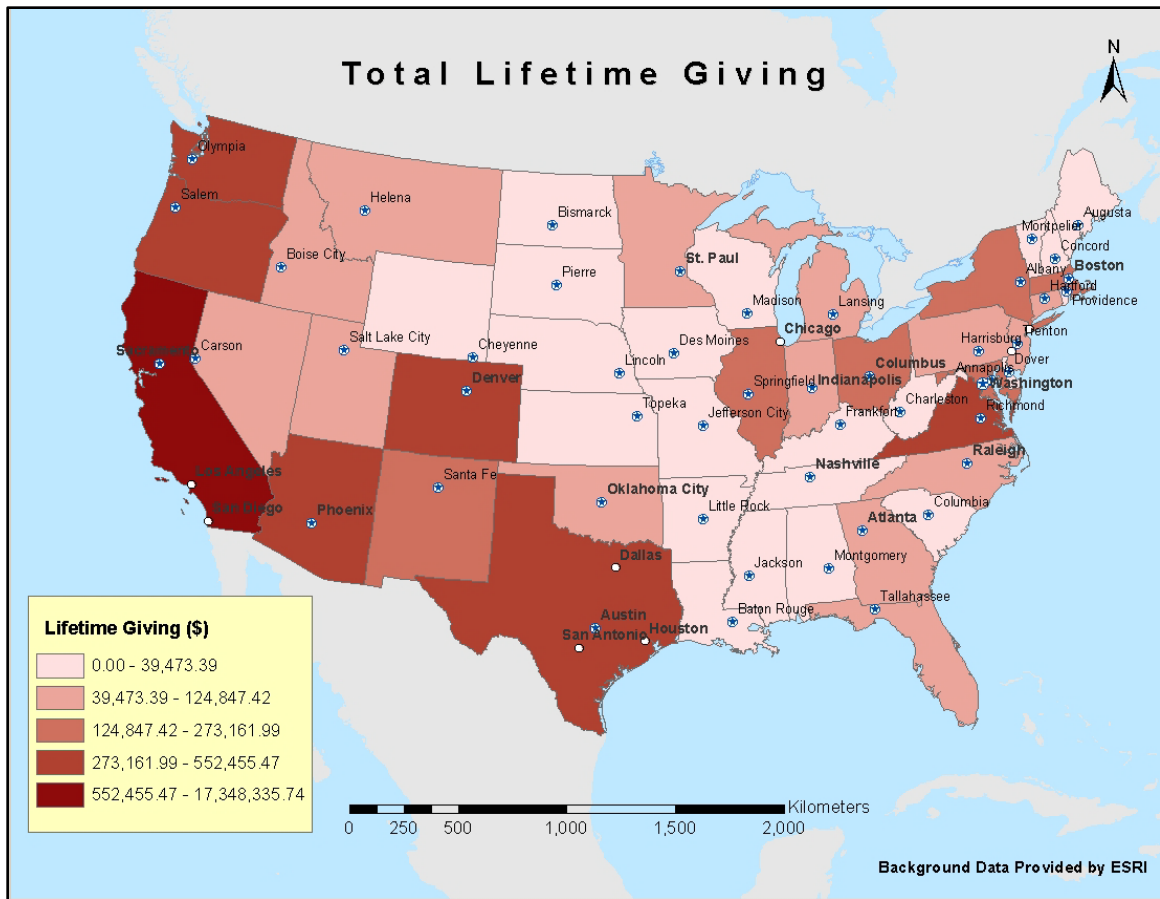


Figure 6.5. Total Giving Map

The average lifetime giving map (Figure 6.6) demonstrates an alternate view of the donation patterns among alumni. Although the majority of donors reside in or near California, the average giving of alumni is highest in central and east coast states. Of the top 10 averages, six are east coast states. There is no west coast state among the highest in average giving, as is shown in Table 6.4.

Table 6.4. Average Lifetime Giving Top 10

State	Avg. Lifetime Giving (\$)	Alumni	Lifetime Giving (\$)	Percentage That Donate (%)
OH	1,840.46	124	228,217.64	43
NM	1,821.08	150	273,161.99	47
DC	1,697.73	49	831,88.65	53
VA	1,631.89	277	452,033.02	44
OK	1,483.74	53	78,638.04	47
NJ	1,356.39	91	123,431.21	51
MA	1,257.77	133	167,283.15	52
RI	1,132.00	5	5,662.00	20
MD	1,002.57	141	141,362.41	45
MT	961.72	75	7,2129.01	41

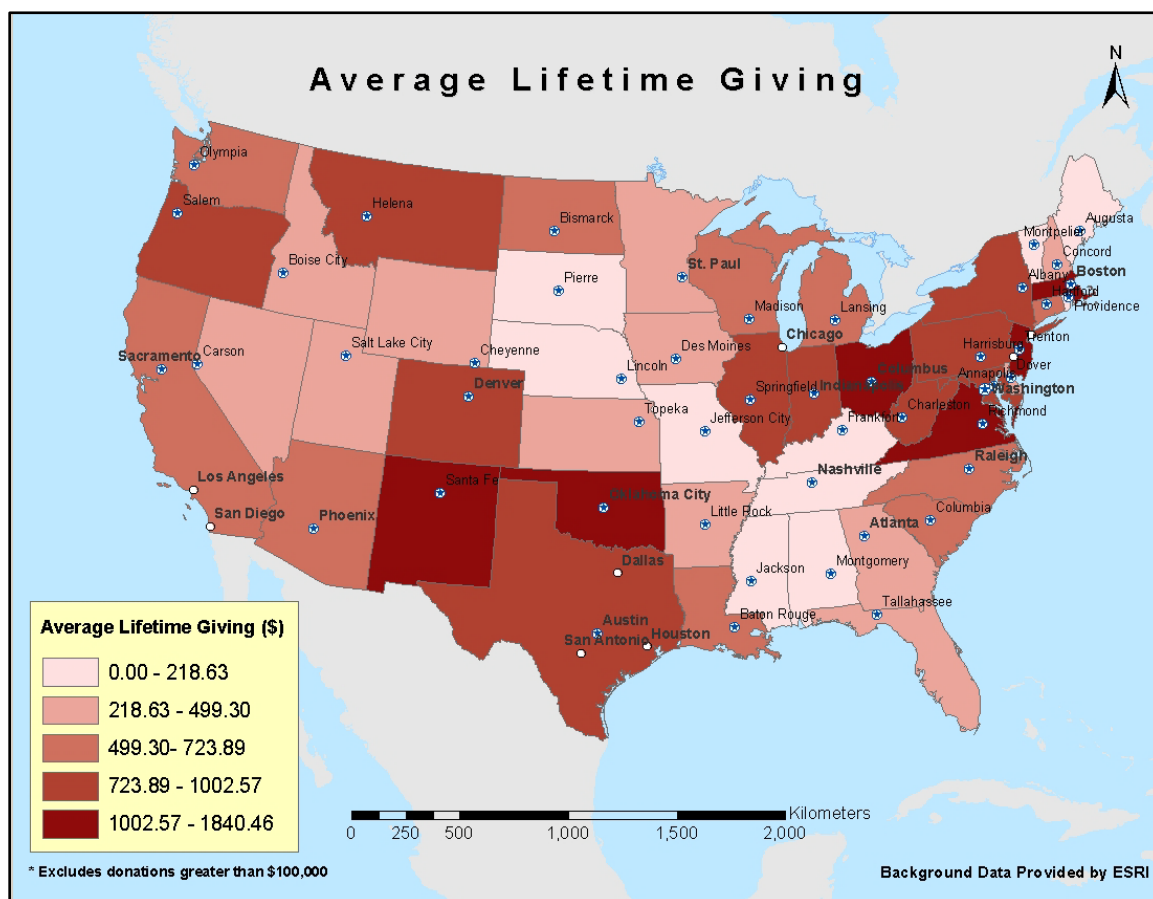


Figure 6.6. Average Giving Map

The final map identifies the top 20 zip codes in terms of lifetime giving within the alumni dataset. Mapping at an increasingly precise scale was helpful for identifying clusters of high donation. Figure 6.7 shows each of the top 20 zip codes are within Southern California, and range from \$143,272 to \$1,667,564. Table 6.5 includes various descriptive statistics for each zip code, as well as Greek member's population percentage, and average graduation time.

Table 6.5 Top 20 Zip Codes Total Lifetime Giving

Zip	State	Total Lifetime Giving (\$)	Total Alumni	Avg. Lifetime Giving (\$)	Percent Greek Members (%)	Avg. Grad. Time (Years)	Percentage that Donate (%)
92373	CA	1,667,564.73	974	1,712.08	20	24	53
92374	CA	732,526.30	676	1,083.62	15	17	41
92660	CA	350,500.16	86	4,075.58	35	30	51
92886	CA	265,325.67	112	2,368.98	15	21	35
91011	CA	239,525.38	45	5,322.79	22	29	60
92024	CA	230,133.43	73	3,152.51	18	23	31
92346	CA	221,057.98	372	594.24	13	16	32
91361	CA	210,598.80	32	6,581.21	28	30	34
92122	CA	206,722.95	42	4,921.98	12	20	40
92651	CA	206,523.12	56	3,687.91	25	33	53
92399	CA	194,588.26	496	392.32	11	16	32
90803	CA	191,297.58	73	2,620.51	25	24	41
93460	CA	186,405.07	9	20,711.67	44	39	66
91107	CA	167,458.70	49	3,417.52	16	28	45
92625	CA	161,671.14	36	4,490.87	39	31	58
92506	CA	158,352.69	252	628.38	13	21	30
92270	CA	156,358.49	37	4,225.91	27	29	49
92313	CA	154,472.13	72	2,145.45	6	19	32
90740	CA	147,151.74	47	3,130.89	23	29	51
91387	CA	143,272.14	16	8,954.51	25	18	37

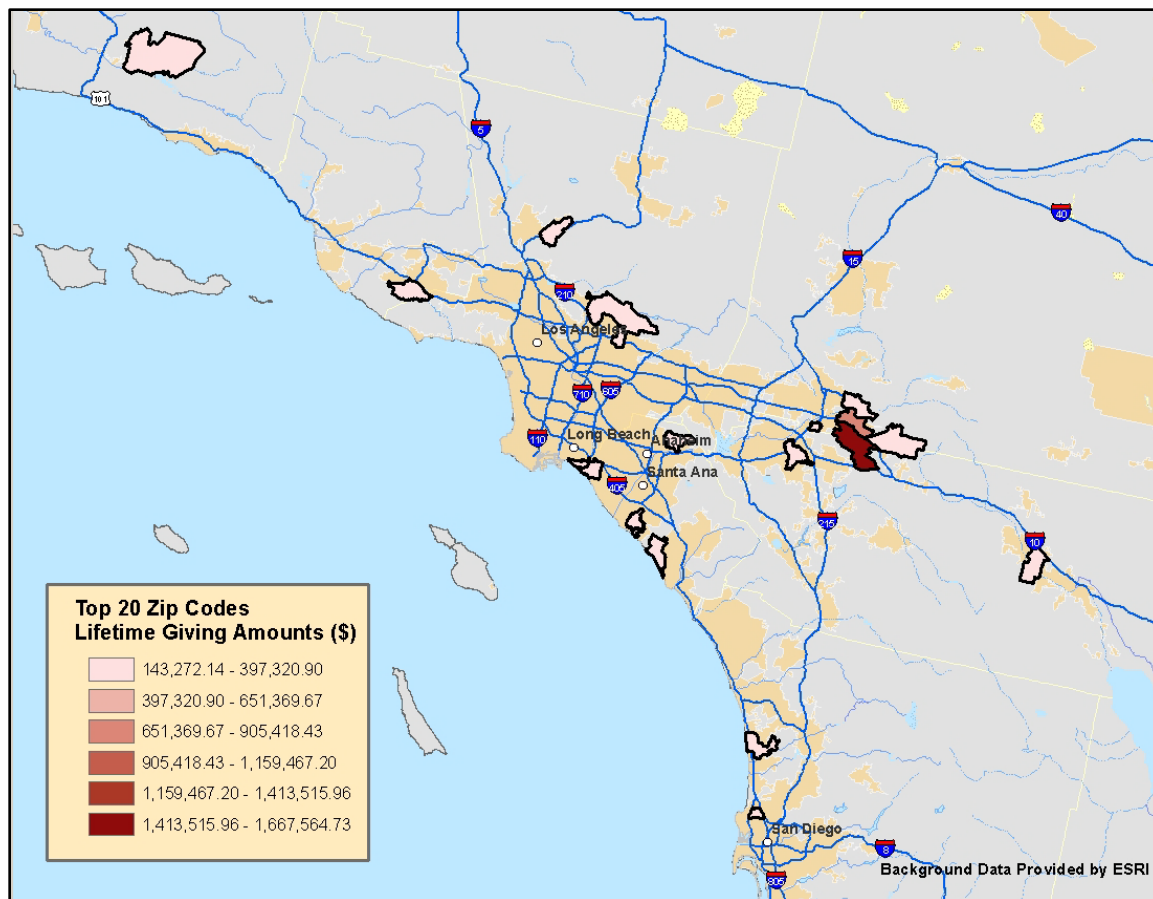


Figure 6.7. Top 20 Zip Codes Map

From the above choropleth maps, it can be seen that distribution of alumni and donations are uneven across the entire county, and the pattern also varies depending on whether the actual totals or the normalized values are considered. For marketing purpose, the maps of totals will tell university staff where the concentration of the alumni and their donations are located. The west coast states are potential areas to target if the university is interested in maximizing the total donations. However, the normalized maps (including percentage who donate and average giving) show the area where people are more likely to donate and on average donate more. In the case of maximizing the benefits of marketing, the east coast states should be focused on.

6.2.2 Does Distance Matter?

According to the literature from Vest (2006), distance from a university may have an impact on the decision to donate. It was expected that as the distance from a university increases, the likelihood of a donation may decrease. A preliminary analysis was conducted regarding this expectation, the results of which are shown on four interrelated maps contained within two figures (Figures 6.8 and 6.9).

Figure 6.8 shows how the total number of alumni changes at different distance ranges across the United States. Over 78% of total alumni live within 800 kilometers or less of Redlands. As the distance increases from Redlands, fewer alumni are found. Less than 1200 total alumni are located greater than 3200 km from the University of Redlands. The total lifetime giving is expectedly highest in and around Redlands, simply because of the distribution of alumni (illustrated in the Figure 6.5). The region within 0 to 100km of Redlands accounts for nearly 45% of the total lifetime giving, which is illustrated in second map within Figure 6.8. The low total alumni and total lifetime giving values between the 200km and 400 km region may be a product of undesirable living conditions within this arid region. Although the eastern region (the area greater than 3200km) has just over half the number of total alumni of the 3200km region, the eastern region still has a greater total lifetime giving than its adjacent region to the west.

Furthermore, although California and the west coast have the highest total of university alumni, the percentage at which they donate is the lowest in the country ranging from 26% to 40% in Southern California. In fact, the percentages at which alumni donate increase as the distance from Redlands increases. The first map within Figure 6.9 shows that states located on the eastern coast of the United States (region greater than 3200km) donated at the second highest percentage with a 44% donation rate. The final map (Figure 6.9) depicts the average lifetime giving of alumni across various distances from Redlands. While the highest lifetime giving totals are centered around Redlands, the average giving of alumni within close proximity to the university is on the lower portion of the scale. The two lowest average lifetime giving amounts are located within the 50km and 200km distance ranges of Redlands. The region of 800km has the highest average giving amount; however, it's noteworthy that the region on the east coast has the second highest average giving amount.

Based on the initial mapping analysis illustrated in Figures 6.8 and 6.9, the expectation that alumni may donate less if their distance from the university is great seems to be false. This finding indicates that the proximity to an alumnus' university does not necessarily increase the loyalty of the individual. The loyalty of an alumnus may actually grow when the individual is farther from the university as shown by the highest average donation at the distance range 800km. This may be due to nostalgic feelings encountered after departing Redlands.

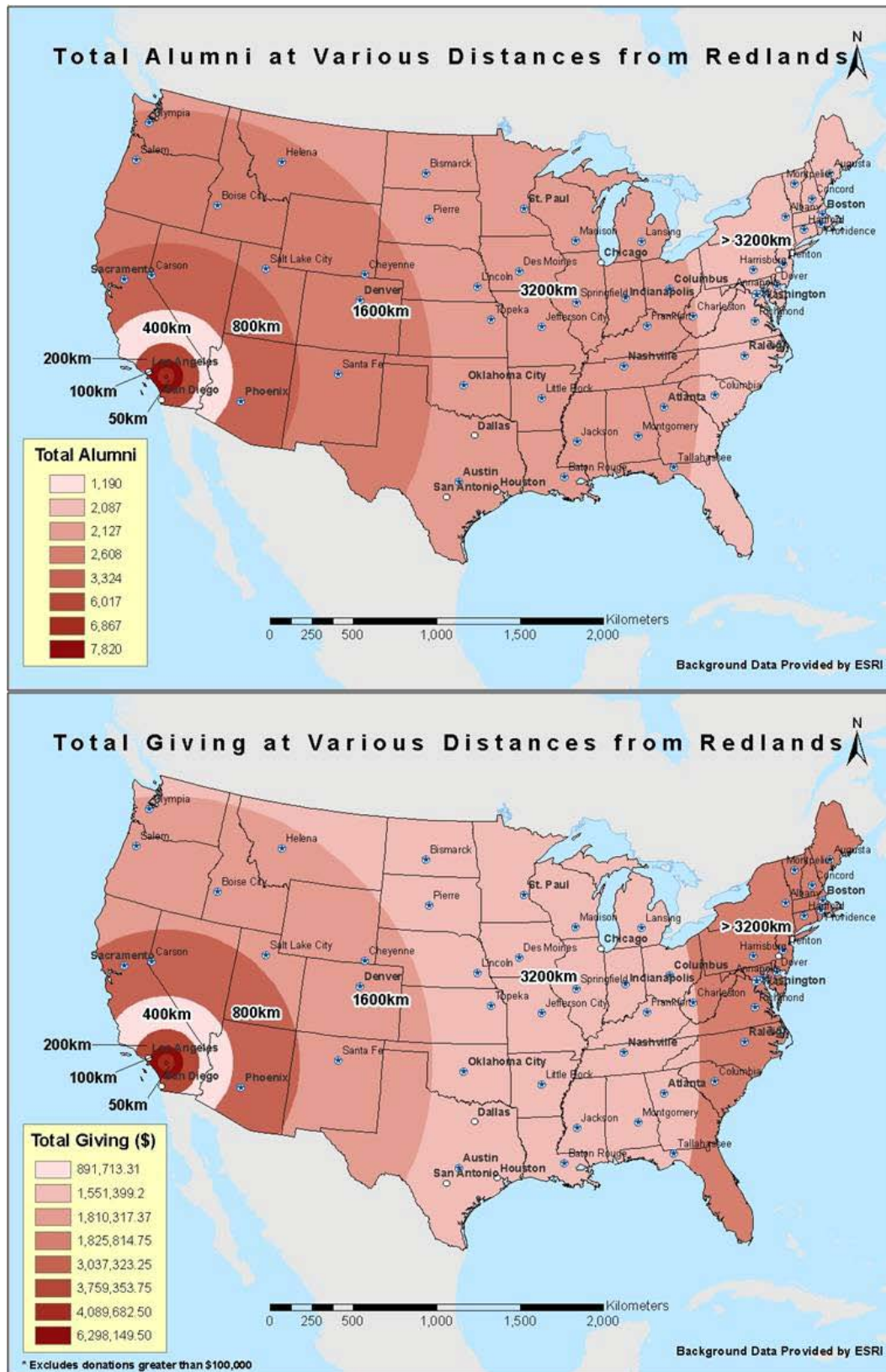


Figure 6.8. Total Alumni and Total Lifetime Giving Distance Map

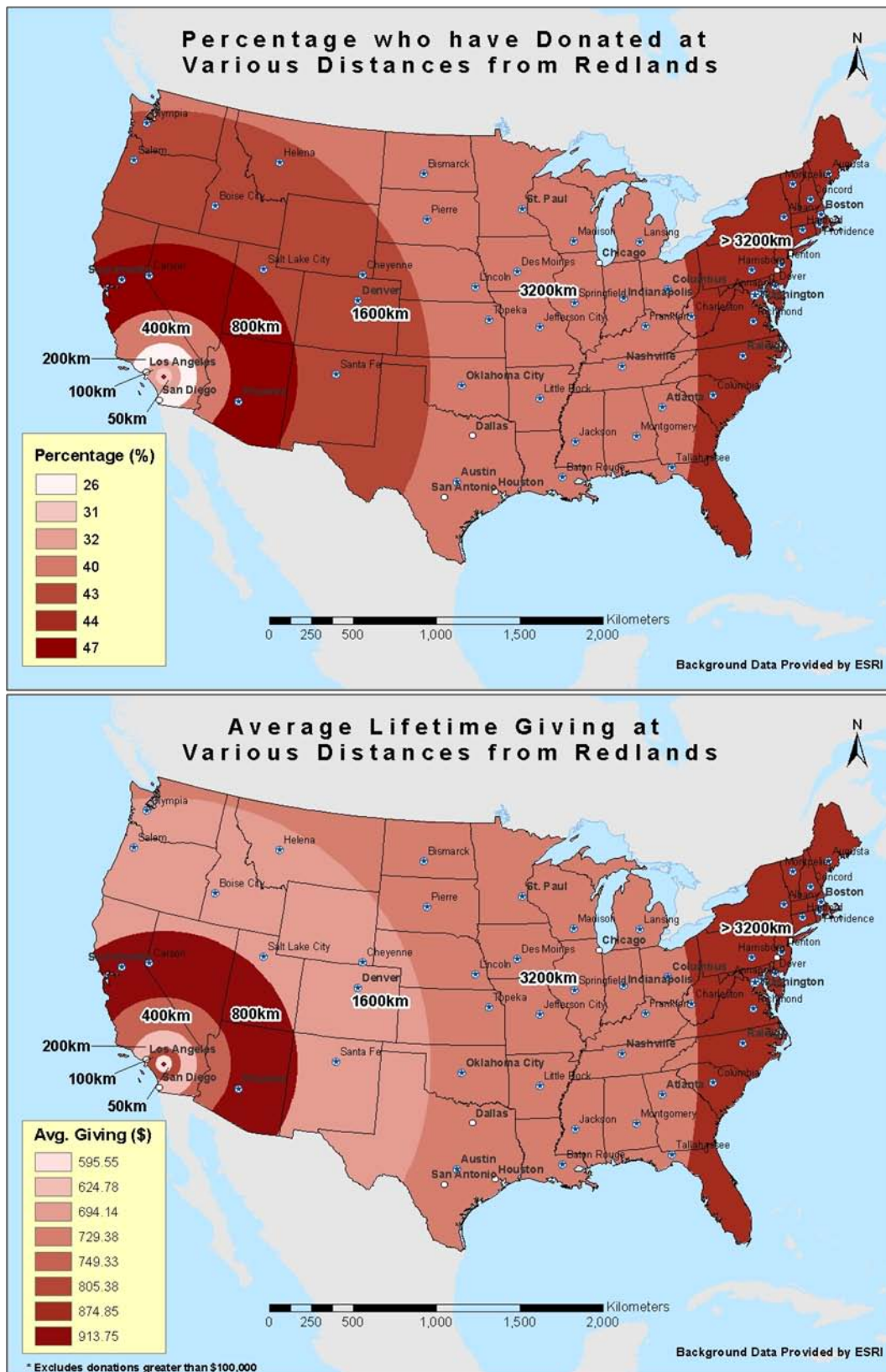


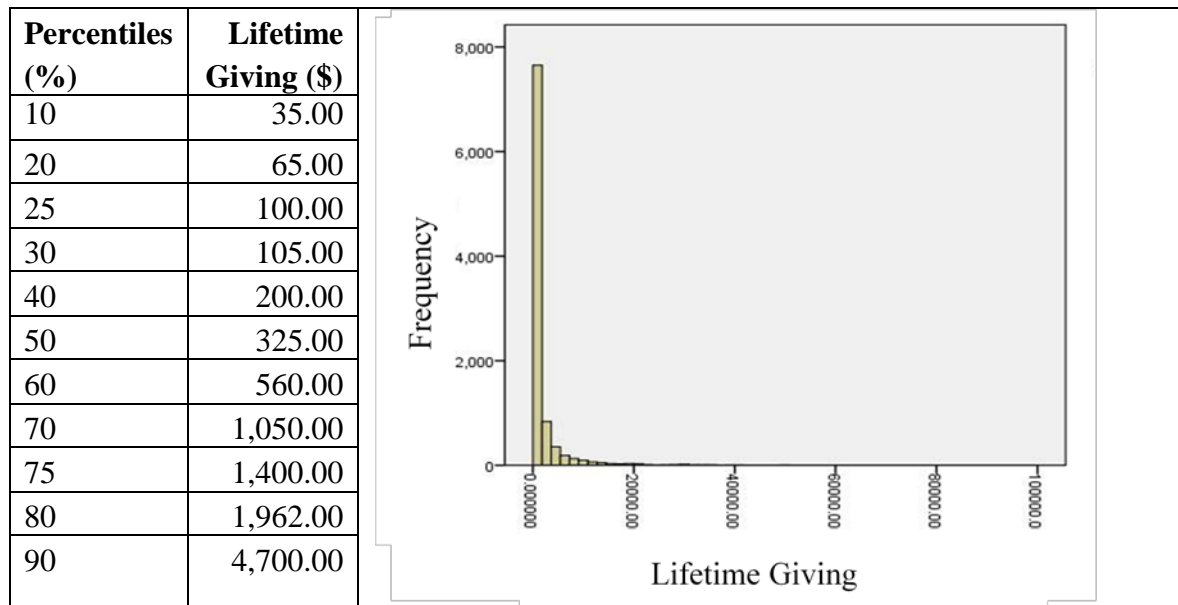
Figure 6.9. Percentage that Donate and Avg. Lifetime Giving Distance Map

6.3 Descriptive Statistics

Descriptive statistics provide the summary measures to describe a data set. Before conducting inferential statistical analysis, it is necessary to generate a hypothesis and examine the distribution of the data set. Therefore, descriptive statistical analysis is often the first step; this section presents the descriptive statistics for the alumni and donation data.

By dividing the data into 10 equal proportions, the percentiles table (Table 6.6) provides a look at the distribution of donations. Among the 9697 donating alumni used in this statistical analysis, 90% of alumni donated \$4,700 or less. The histogram in Table 6.6 indicates that data is extremely skewed. The analysis conducted in Section 6.4 discusses the problems encountered while working with such a heavily skewed data set.

Table 6.6. Percentiles Summary of Donating Alumni Lifetime Giving



The Greek membership summary table (Table 6.7) provides summary statistics of the 15 fraternities and sororities sanctioned at the University of Redlands (the meaning of these group abbreviations can be found in the list of acronyms in the report). Only one-third of the Greek households have donated at a less than 50%, while the overall average remains high at 72.8%, indicating that Greek membership might be a significant variable affecting donation. The relationship between a Greek membership and the likelihood of a donation is further analyzed in Section 6.4.

Table 6.7. Greek Membership Summary

Group	Lifetime Giving Sum (\$)	Total	Avg. Lifetime Giving (\$)	Percent That Donate (%)
DKP	3,129,813.81	1034	3,026.90	78
KSS	2,268,702.06	698	3,250.29	76
PIC	1,852,114.63	530	3,494.56	77
ATP	1,772,071.38	957	1,851.69	74
CSC	1,230,187.78	518	2,374.88	73
ASP	1,155,201.77	562	2,055.52	79
BLM	449,197.97	385	1,166.75	65
AXO	348,093.70	283	1,230.01	70
SKA	51,906.18	109	476.20	39
PST	32,930.89	148	222.51	41
ACD	8,175.15	7	1,167.88	86
PBS	4,384.83	28	156.60	36
ZPB	2,250.00	5	450.00	40
CRP	1,295.20	29	44.66	41
APA	618.00	4	154.50	50

An initial look at individual alumni and the number of years since gradation revealed a positive correlation between the increase in years since graduation and the percentage at which alumni donate. To further analyze this relationship, the variable was grouped into eight categories using equal intervals within the data. Depicted in Table 6.8 below, the eight categories reveal that alumni graduating between 1 to 10 years ago represent the highest total with 29.8% of the alumni population, yet this group donates at the lowest percentage of the eight categories (14%). As the number of years since graduation increases, so does the percentage at which alumni donate, and the average amount they give. Therefore, the number of years could also significantly affect donation behavior.

Table 6.8. Years since Graduation Summary

Years Since Graduation	Total	Lifetime Giving (\$)	Avg. Lifetime Giving (\$)	Percent that Donate (%)
1 to 10	9304	193,150.37	20.76	14
11 to 20	7638	1,260,036.43	164.97	22
21 to 30	5423	2,605,039.84	480.37	37
31 to 40	4113	4,664,100.37	1,133.99	56
41 to 50	2450	7,643,386.97	3,119.75	77
51 to 60	1570	4,083,820.94	2,601.16	77
61 to 70	557	1,753,182.68	3,147.55	80
71 to 81	98	454,693.24	4,639.73	84

The final portion of the descriptive statistical analysis explores the distribution of alumni among various majors. The majors available at the university have been categorized into natural science, social science, humanities, education, business/economics/professional, interdisciplinary, and uncategorized. The top three disciplines for the average lifetime giving are natural science, social science, and humanities as shown in Table 6.9 and Figure 6.10. These three disciplines also have the highest percentage of alumni who donated, the highest percentages of Greek population, and relatively high average graduation years. In contrast, business /economic/professional degrees have graduated the most alumni; however they donate at the second lowest amount and percentage. When the disciplines are separated between bachelor's degree (Table 6.10) and graduate degree (Table 6.11), the results are similar. Three of the top four disciplines (social science, humanities, and education) remain the same. These three disciplines may have high average lifetime giving amounts because of their active participation in a fraternity or sorority accompanied with higher than average values for years since graduation.

Examining the low average donation amounts of the most populous group (business, economics, and professional) further, it's worth noting that the alumni of this discipline participates the least in Greek memberships, and is among the lowest in average years since graduation. Within this discipline, 29% of degrees awarded have been graduate level degrees. Further, small standard deviations of the average lifetime giving of this group also indicate the relatively small variation in donations among the alumni of this discipline. Therefore, although one may speculate that alumni graduating with a degree from a business discipline may have a greater income and thereby they may donate more, it's clear that this group's loyalty may lie elsewhere.

Table 6.9 All Disciplines Summary Table (Bachelor's and Master's Degree)

Discipline	Avg. Lifetime Giving (\$)	Std. Deviation (\$)	Total	Total Lifetime Giving (\$)	Percent that Donate (%)	Percent Greek (%)	Avg. Grad Time (Yr)
Natural Sciences	1,773.92	92.53	1,824	3,235,633.65	57	27	28
Social Sciences	1,499.13	159.88	2,967	4,447,924.76	59	37	28
Humanities	1,101.66	115.59	4,124	4,543,236.39	56	28	30
Education	992.11	101.45	1,974	1,958,432.15	43	23	33
Uncategorized	895.71	137.14	1,848	1,655,279.48	51	19	29
Bus/Econ/Prof	392.96	92.53	17,159	6,742,756.18	21	8	18
Interdisciplinary	58.99	25.34	1,257	74,150.23	23	14	8

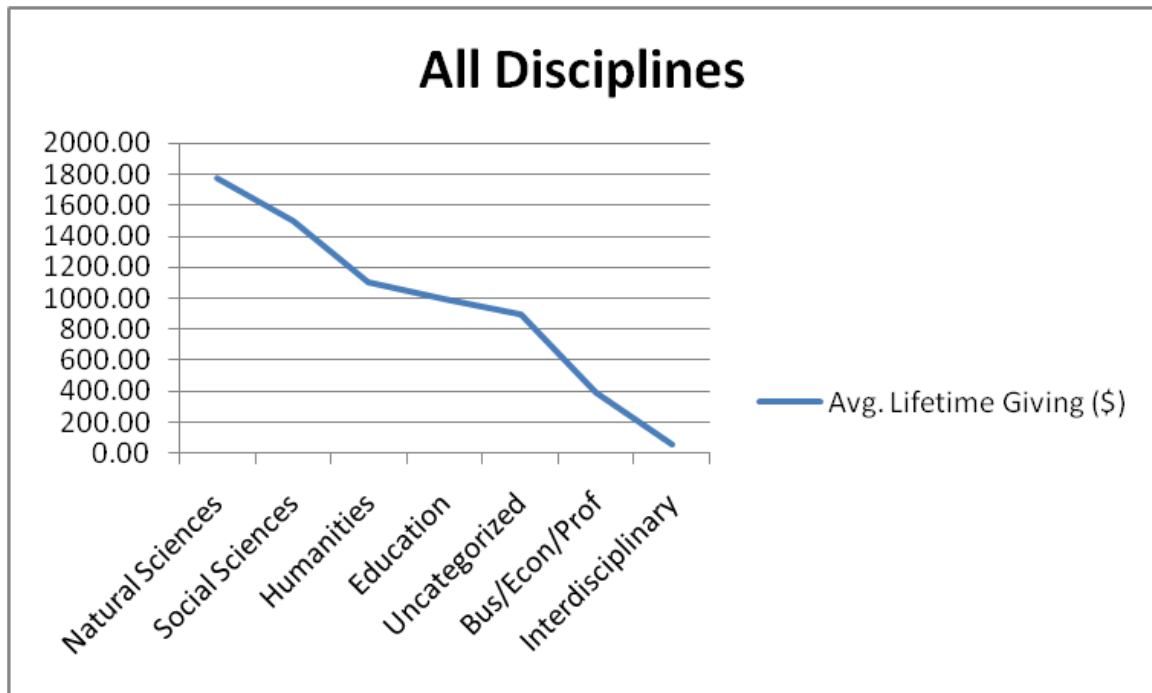


Figure 6.10. All Disciplines Summary Graph

Table 6.10. Summary Table for Alumni with a Bachelor's Degree

Discipline	Avg. Lifetime Giving (\$)	Std. Deviation (\$)	Total	Lifetime Giving (\$)	Percent that Donate (%)	Percent Greek (%)	Avg. Grad Time (Yr)
Education	2,176.00	124.42	636	1,383,937.30	77	59	55
Natural Sciences	1,774.88	173.36	1823	3,23,5612.65	57	27	28
Social Sciences	1,498.68	160.27	2951	4,422,625.95	59	37	27
Humanities	1,127.24	117.90	3913	4,410,907.26	57	30	30
Uncategorized	1,045.03	150.32	1522	1,590,542.43	56	22	28
Bus/Econ/Prof	489.40	95.27	12217	5,979,083.72	25	11	19
Interdisciplinary	58.86	25.38	1253	73,755.23	23	13	8

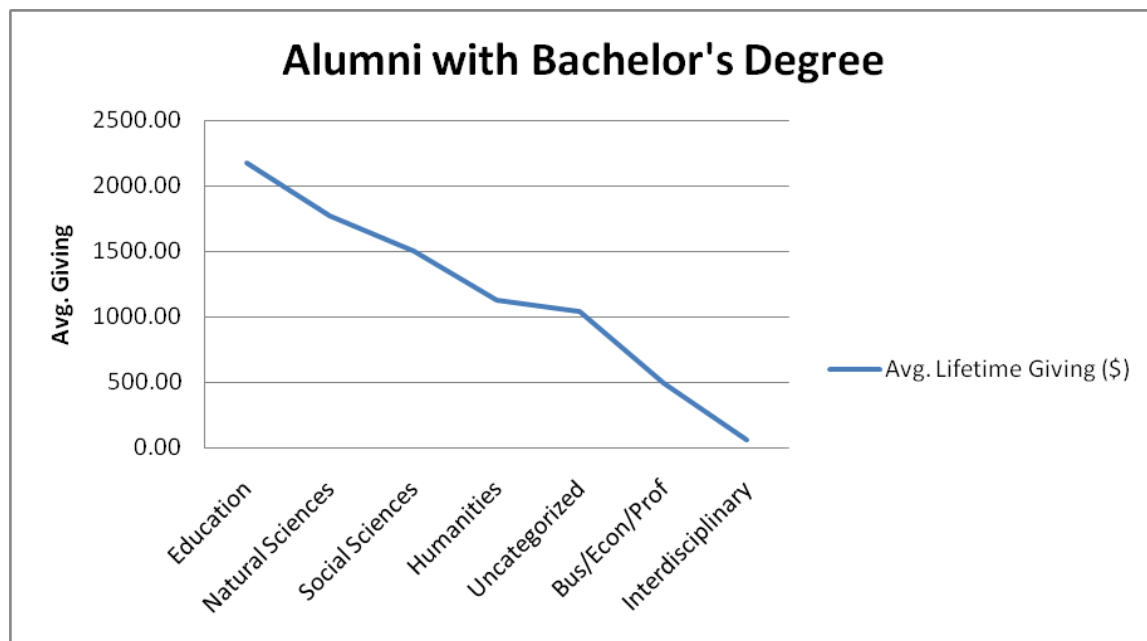


Figure 6.11. Bachelors Degree Summary Graph

Table 6.11 Summary Table for Alumni with a Graduate Degree

Discipline	Avg. Lifetime Giving (\$)	Std. Deviation (\$)	Total	Lifetime Giving (\$)	Percent that Donate (%)	Percent Greek (%)	Avg. Grad Time (Yr)
Social Sciences	1581.11	51.89	16	25297.81	87	50	51
Humanities	627.15	56.54	211	132329.13	44	6	32
Education	429.37	87.22	1338	574494.85	28	6	22
Uncategorized	198.58	22.36	326	64737.05	26	3	30
Bus/Econ/Prof	154.53	85.19	4942	763672.46	14	2	13
Interdisciplinary	98.75	3.08	4	395.00	100	25	28
Natural Sciences	20.00	0	1	20.00	100	0	44

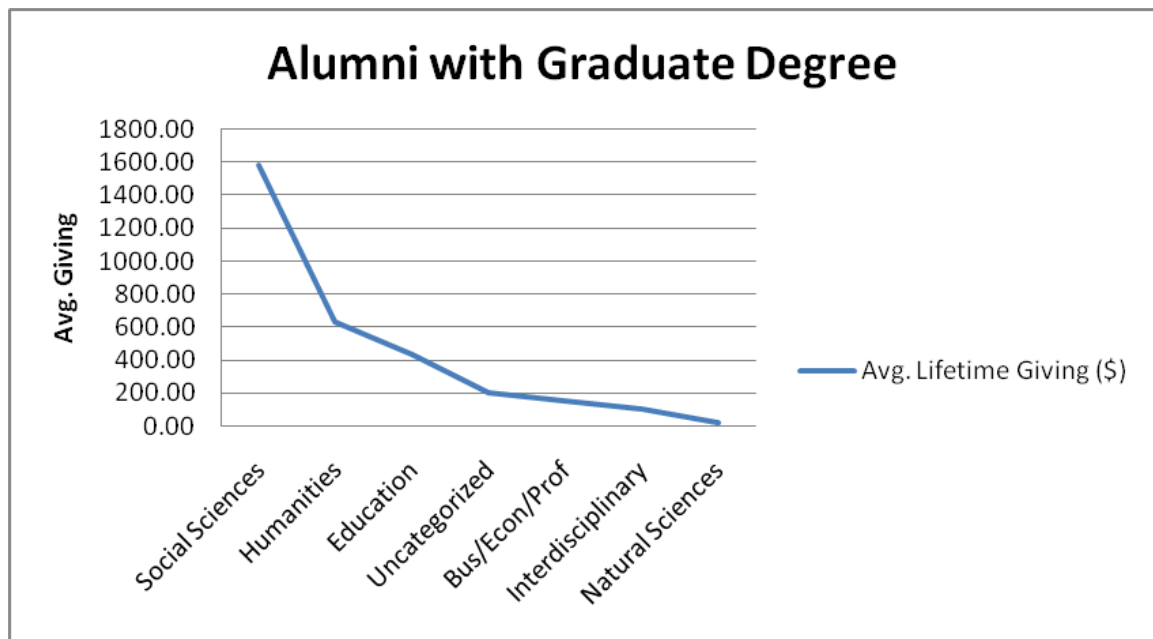


Figure 6.13. Graduate Degree Summary Graph

6.4 Statistical Results

This section discusses analyses results from the four statistical methods mentioned in Section 5.3. The four methods employed in this section sought to provide a clearer profile of a donor; specifically, who is most likely to donate and which variables show who will likely donate more? The explanatory variables used in these models are summarized in Table 6.12. Besides the two variables for the individual donors, another three variables pertaining to the geographic context in which the alumni live were considered. It is expected that the alumni who live in a neighborhood that is relatively wealthy, white-dominated, and that have a high percentage of residence that are well

educated are more likely to donate. Simply, it attempted to test if the geographic indicators have an impact on individual donation behavior. However, it should be noted that the models tested in this study have limitations due to the lack of other individual attributes that might also account for the variations of individual gift giving.

Table 6.12. Explanatory Variables

Variable Names	Description
Greek Membership	The Greek member variable identifies whether alumni have ever participated in any Greek household, with 1 representing yes and 0 representing no.
Graduation Time	The graduation year variable identifies the number of years that have passed since an alumnus was awarded a degree.
Percent Caucasian	The percent of population that are Caucasians in the residential area at the zip code level.
Percent Educated	The percentage of individuals with a bachelor's degree above the age of 25 at the zip code level.
Median HH Income	The median household income at the zip code level.

6.4.1 Who are More Likely to Donate?

The independent samples T-Test compared the mean value of continuous variables against one grouping variable. The four continuous variables studied were percentage of Caucasians, percentage of individuals with a bachelor's degree, years since graduation, and median household income. The grouping variable used was donation, which is a binary variable with either a zero (have not donated), or a one (have donated). Table 6.13 presents the resulting table showing the difference in means among the four variables.

The mean value from each continuous variable is greater in the donation group in all four cases. The mean percentage of individuals with a bachelor's degree is higher among donating alumni (15.5%) than among non-donating alumni (13.9%). The mean percentage of Caucasian individuals is higher among donating alumni (69.8%), than non-donating alumni (66.1%). The mean value of median household income is higher among donating alumni (\$73753) than non-donating alumni (\$71404). Finally, the mean value of years since graduation among donating alumni is twice that of non-donating alumni. More importantly, the T-Test result shows that the difference between the group means are significant across all four variables, confirming that alumni who donated are more likely to be those who graduated longer and live in a distinctive neighborhood.

Although each variable was shown to be significant, it's important to acknowledge the affect that the large sample size (n) had on the T-Test. In this particular case, the large sample size leads to the small standard deviations of sample means in the T-Test and therefore it is easy to obtain a significant result. Even though the mean difference is not large for some variables (e.g. percent Caucasian of the two groups are 66% and 69%), the test yields significant results.

Table 6.13. Independent Samples T-Test Results

	Donation Group	Mean	F	Sig.
Percent Educated	0 no	.139450	38.119	.000
	1 yes	.155522		
Percent Caucasian	0 no	.661385	9.418	.000
	1 yes	.698751		
Median HH Income	0 no	71,404.62	33.944	.000
	1 yes	73,753.79		
Graduation Time	0 no	16.69	78.152	.000
	1 yes	32.30		

Expanding upon who donates, the Greek membership variable was further examined. Using the two sample difference in proportions test, the proportion of Greek members within the donate group (34.4%) was compared to the proportion of Greek members within the non-donating group (6.9%). The proportion test result showed a p value close to zero which indicated that the proportion of Greek members within the donating alumni group is significantly higher than the proportion within the non-donating alumni group. Therefore, it seems that donating alumni group shows a distinctive characteristic that the non-donating group lacks.

6.4.2 Who are More Likely to Donate More?

To answer the second question, the study focused on the alumni who have donated (the donating group). In this situation, linear regression was used to investigate the relationship between the average lifetime giving and the five explanatory variables. The ordinary least squares (OLS) procedure within ArcMap was executed and the results are shown in Table 6.14. Three of four independent variables were significant (Gradtime, MedHHinc, PercEdu). The F statistic also yielded significant result indicating at least one variable within the model was significant. The extremely small R-squared value indicates only about 2.4% of total variance was explained therefore residual variability remained high, which suggested poor model performance.

Table 6.14. OLS Results

Dependent Variable: Average Lifetime Giving

Joint F Statistic = 47.82

R-square = .024

Independent Variables	Coefficient	Std. Error	Sig.
Percent Educated	168.8198	39.4638	.0000
Percent Caucasian	6.1616	11.8106	.6018
Med. HH Income	.0003	.00009	.0002
Graduation Time	.8527	.1180	.0000

There are multiple reasons that this first phase regression analysis was inconclusive, including both lack of independent variables and the degree to which the data was skewed. The alumni data was lacking in comprehensive variables to develop a more suitable regression model. Information such as alumni gender, income, and number of children could have increased the fitness of the OLS model. Regression analysis was also made difficult by the extreme degree to which the alumni donation patterns were skewed. Ninety percent of alumni among the donating group donated less than \$4700 (lifetime giving).

Thus, the one-way ANOVA (analysis of variance) test was used as an alternative to OLS. Using average giving as a factor variable, the mean of each explanatory variable were compared with respect to categorical values of average giving. These groupings were as follows:

- \$0-\$50 = 1
- \$50-\$100 = 2
- \$100-\$200 = 3
- \$200-\$500 = 4
- \$500 and above = 5

In addition to the results table found in Table 6.15, the procedure created a series of mean plots shown in Figure 6.14. The results from this test suggest the alumni who fall in different categories of donation amounts have significant characteristics in terms of where they live, the number of years since they graduated, and whether they participated in any Greek household. In addition, each mean plot in Figure 6.14 shows a generally positive correlation between average giving amounts and the increase in each explanatory variable's value. Combined the mean plots with the statistical testing results, it seems that alumni who donate more are those who graduate longer ago and live in a wealthy, white dominated and well educated neighborhood. However, this does not apply to all the situations. For example, the neighborhood of the alumni whose average donation is

between 200 to 500 dollars actually has a lower percent white population than that of the alumni who donated between 100-200 dollars. Another issue with this analysis is grouping schema. Because the grouping cut values used in this study are not natural, slightly different results may be possible for other grouping schema.

Table 6.15. One-Way ANOVA Results

		Sum of Squares	F	Sig.
PercWhite	Between Groups	1.138	9.41	.000
	Within Groups	292.56		
	Total	293.69		
PercEdu	Between Groups	.70	38.11	.000
	Within Groups	44.74		
	Total	45.44		
MedHHInc	Between Groups	1.152E1	33.94	.000
	Within Groups	8.219E1		
	Total	8.334E12		
GradTime	Between Groups	88,268.47	78.15	.000
	Within Groups	2,734,372.93		
	Total	2,822,641.41		

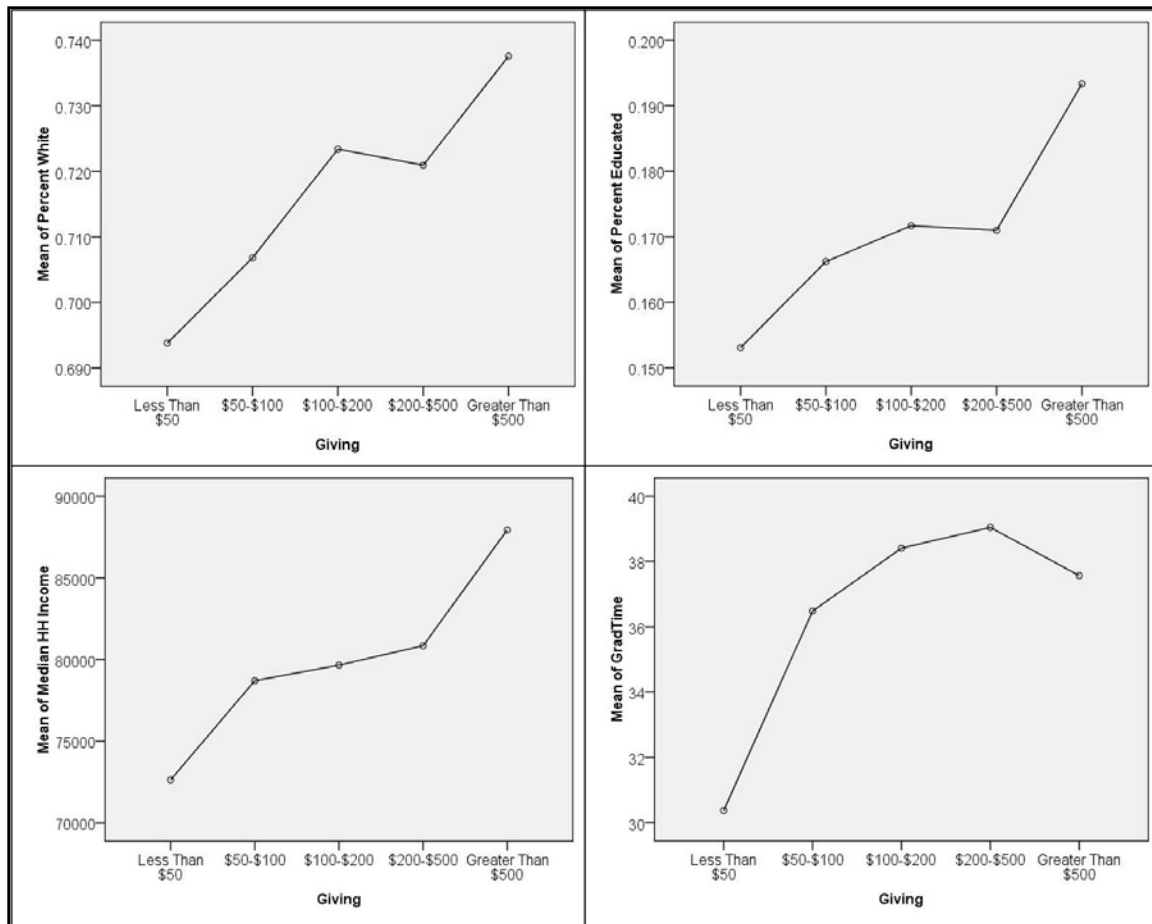


Figure 6.14. ANOVA Mean Plots

6.4.3 Mapping Targeted Potential Donors in Southern California

Working with the donor profile developed from the previous sections, it is possible to exploit the variables to identify those who have this donor profile but have not donated yet. This portion of the analysis will focus on Southern California as an example of how the office staff may conduct marketing in the future. Figure 6.15 illustrates the sample study region in which potential donors were to be identified.

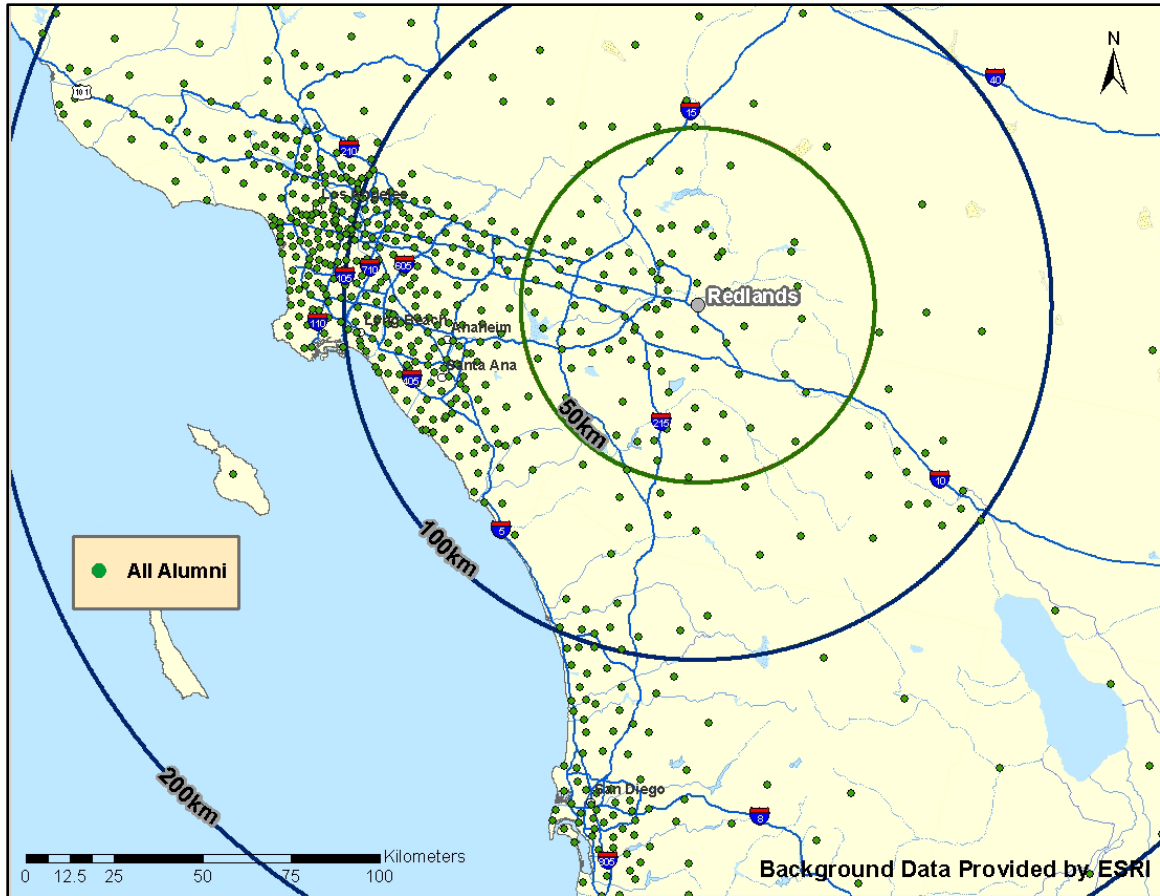


Figure 6.15. All Alumni Southern California

Three methods were utilized within ArcMap to identify these potential donors. First the mean value of each neighborhood statistic for *PercWhite*, *PercEdu*, and *MedHHinc* were identified within the study area by using the statistic summarization tool within the attribute table of the alumni point data. These values were used as a threshold for eliminating alumni whose neighborhoods were below these mean values. A definition query isolated only donors with the following attributes:

- Have not donated
- Greek members
- Greater than 14% of population within zip code has a bachelor's degree
- Percentage of Caucasians within the zip code was greater than 62%
- Median Household income was greater than \$82,000
- Years since graduation was greater than 10 years

The remaining 110 alumni are shown in Figure 6.16. This map shows all alumni who meet the requirements listed above; the alumni points on the map are symbolized by

number of years since graduation.

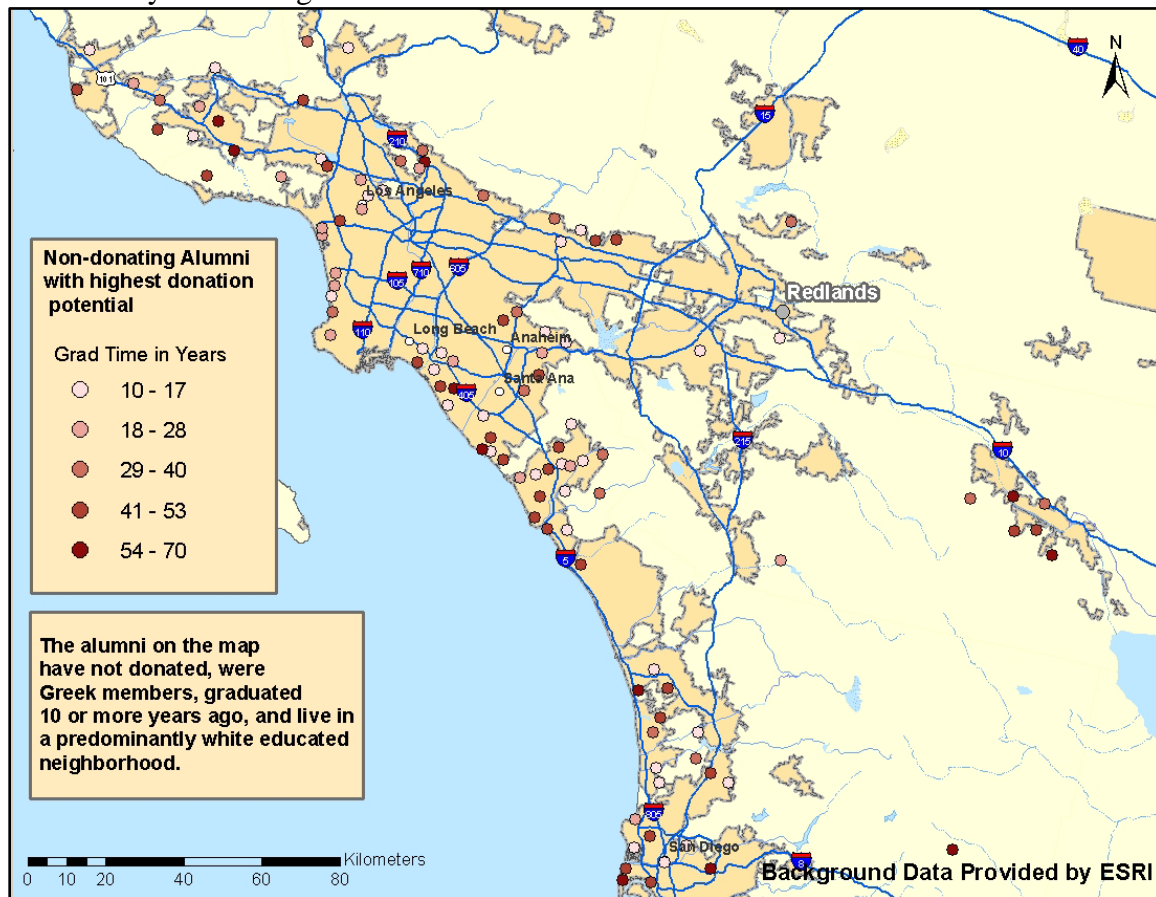


Figure 6.16. Potential Donors Southern California

While all 110 donors in Figure 6.16 could potentially be targeted; additional spatial analysis (kernel density) identified high potential alumni who overlapped because of the scale at which the data was geocoded. The first map in Figure 6.17 shows the areas in which there were concentrations of alumni. The second map in Figure 6.17 isolated one of these areas, and identified four alumni who have a high likelihood of donating based on the analysis conducted in Sections 6.4.1 and 6.4.2. Finally, the four alumni with potential to donate are presented in Table 6.16.

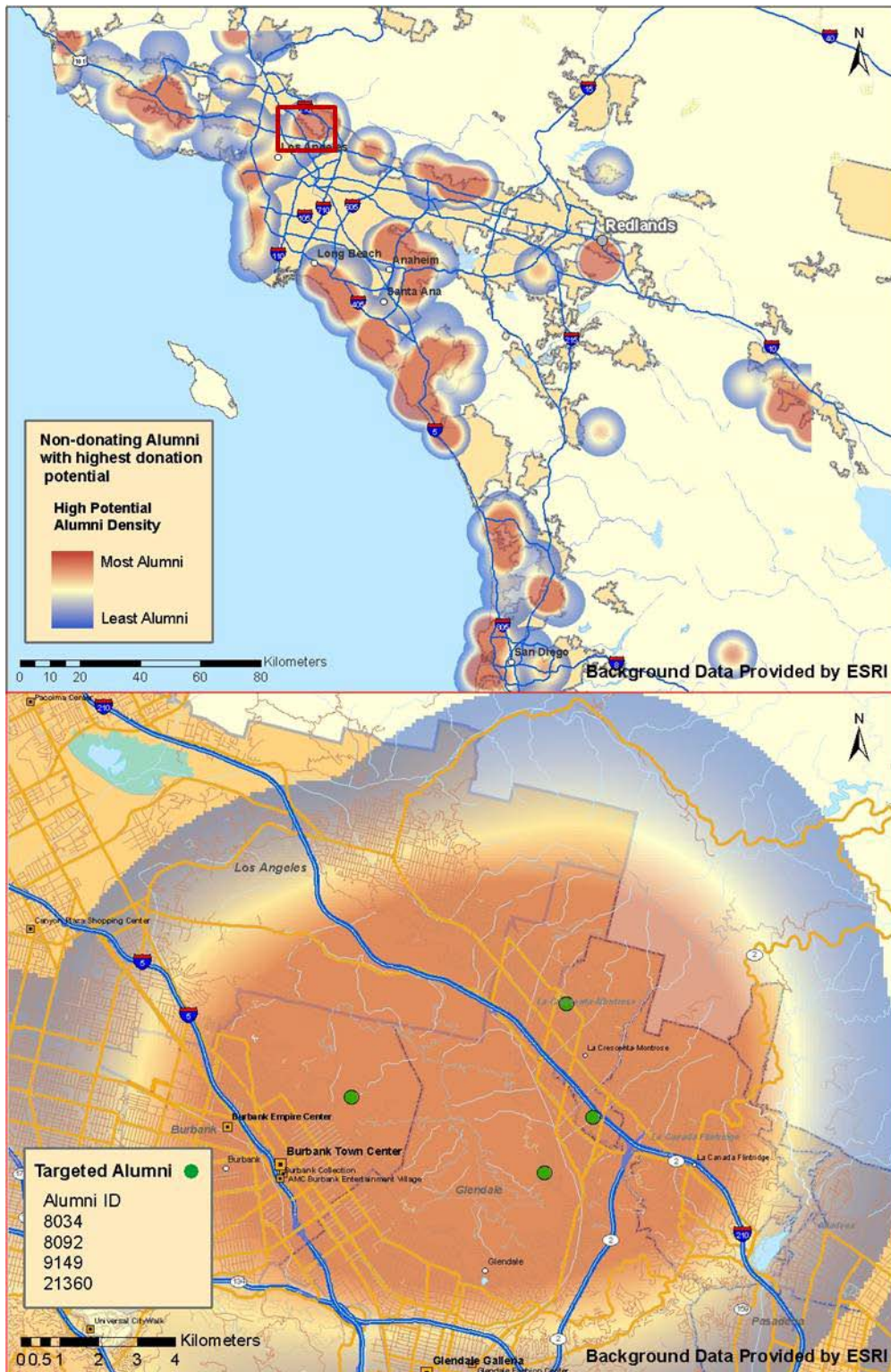


Figure 6.17. Alumni Density Map and Targeted Alumni in S. California

Table 6.16 Example of Four Targeted Alumni

Alumni ID	Grad. Time	Greek	City	Zip
9149	13	PST	Glendale	91208
8034	36	DKP	La Crescenta	91214
8092	33	ATP	La Crescenta	91214
21360	18	BLM	Glendale	91208

6.5 Summary

Chapter Six provided commentary on the successful functionality of the Alumni Data Explorer. While only a prototype, the ADE demonstrated the possibilities available to the client concerning database management, and the usefulness of exploring alumni data within a mapping environment. The chapter supplied a discussion on a series of maps generated using spatial join techniques within ArcMap. These maps gave insight into the donation patterns on a national scale, and introduced the graphical depiction of the relationship between distance to the university and donation patterns. Various statistical methods were researched to study the donation patterns of alumni; four were employed in this project. From the data available, it is clear that membership in a Greek household and the number of years since graduation is the most telling variables among the attributes within the alumni data. Using these variables and other neighborhood statistics, alumni with a high potential to donate can be identified via mapping and spatial analysis. With the augmenting of additional variables within the data, such as alumni gender or income, accompanied with the use of more socio-economic variables from ESRI's 2008 demographic dataset, a more conclusive picture may be generated.

Chapter 7 – Conclusions and Future Work

The goal of this project was to help the University of Redlands Development Office analyze the alumni network and therefore aid in the allocation of the client's limited resources when targeting potential or existing donors. This objective was supplemented with a normalized relational database to better serve the client's efforts to maintain data. In addition to the database, a prototype data entry application was introduced, along with the concept of exploring alumni data in a GIS client. In summary, the critical success factors were:

1. Provide a normalized relational database schema and populated prototype database that is conducive for use with GIS technology.
2. Provide a prototype custom application to demonstrate the usefulness of a relational database, and the functionality of the free client known as ArcGIS Explorer.
3. Provide the client with analyses of the University of Redlands donor network by utilizing various spatial and statistical analysis methods. Depict these findings with maps and various other graphics, and provide a thorough write-up of these findings.

Of the three distinct components within this report, two had a list of functional requirements (found in Chapter Three) that could be used to measure the success of the project. The alumni data explorer (ADE) had five mandatory functional requirements that all related to manipulating the data. Each requirement was successful. The functional requirements for the alumni database included providing the development office with an efficient method for organizing data, normalizing this data for optimal performance, and indexing searchable fields within the database. Each of these three functional requirements was successful.

Item three from the above list sought to identify relationships, patterns, and characteristics within the existing donor data that could provide a clearer picture of only who donates, and the factors which most likely influence donation amounts. This process began by utilizing various techniques within ArcMap to create both tables and maps to illustrate the distribution and donation statistics of University of Redlands alumni across space. Through the identification of several variables discussed in Chapter Five and Six, the characteristics that could identify the alumni most likely to donate have become clearer. Subsequent to identifying these important characteristics within the existing data, three additional variables were tested to further improve upon this profile. Employing a method in which neighborhood demographic statistics were used to test the significance of the explanatory variables identified in the data, the validity of the explanatory variables identified as distinctive among donating alumni was increased. The findings were that the alumni who donate and who donate more are those who graduate longer, participated in a Greek household, and live in relatively wealthy, white dominated, and highly educated neighborhoods.

Significant effort was spent creating project objectives, identifying a scope, and developing and imparting a solution. There are several areas in which the project could be expanded upon, including:

- An implementation plan to provide a complete GIS based solution
- Further development of relational database and database application
- Additional statistical analyses using more complex techniques and more variables

A complete implementation plan, would seek to provide a more comprehensive needs assessment pertaining to the client's day to day business practices, and outline an agenda to put into place a position within the office dedicated to the GIS field. Roger Tomlinson (2007) identifies 11 key issues to address when creating an implementation plan:

- Department staffing - Could the development office benefit from a GIS professional?
- Training regime - Could current staff receive GIS training?
- Institutional interaction requirements
- System requirements - What type of software and computer(s) are necessary
- Legal review
- Security issues – Are there issues with protection of this data on a network?
- Existing computing environments - Is the existing network and hardware set-up sufficient?
- Migration strategy - How will data be migrated to data suitable for use within a fully fledged GIS?
- Risk analysis – What are the risk areas when implementing such a system?
- Alternative implementation strategies – Are there additional strategies?
- System procurement

With a GIS professional in place, analysis and mapping could be completed on an ad hoc basis when information about a particular neighborhood or group of alumni is needed. Using methods such as those discussed in Section 6.4.3 could provide the client with spatial analysis on any portion of the alumni network as needed.

Second, transferring the database to a server-based product such as SQL Server may provide additional database resources previously unavailable with the use of Microsoft Access. Expanding upon this solution, further development of a standalone database application could also be a future expansion point for this project. Additional advanced programming incorporating a server based database and further development of a stand-alone application would allow for more users, advanced queries, and an overall increase in capabilities as they pertain to data storage and manipulation.

The final expansion point involves delving into more advanced statistical techniques to further scrutinize the alumni data. The linear regression techniques exercised in this project yielded models that provided incomprehensive results. Through the exploration of the hundreds of additional variables provided within ESRI's tapestry data set, and using additional techniques such as multinomial regression, the profile of a donor would become increasingly clear. In addition to more statistical techniques, the relationship between distance and alumni giving could be examined more thoroughly by looking at the data on a larger scale and implementing or creating a distance decay model.

Works Cited

- Anselin, L. (1995). Local indicators of spatial association-LISA. *Geographical Analysis*, 27(2), 93-115.
- Chandra, N. (2007). *The study of fund raising process*. (Project Report, Xavier Institute of Management & Entrepreneurship), 2007.
- ESRI. (2002). *Geographic information systems in philanthropy* [White Paper]. Retrieved October 7, 2008 from ESRI Library:
http://www.esri.com/library/whitepapers/pdfs/gis_in_philanthropy.pdf
- Frumkin, P. (2006). *Strategic giving: The art and science of philanthropy*. Chicago: Univeristy of Chicago Press.
- Getis, A., and J.K. Ord. (1992). The analysis of spatial association by use of distance statistics. *Geographical Analysis*, 24, 189-206.
- Gimpel, J. "Prospecting for Campaign Gold: Predicting the Spatial Distribution of Campaign Contributions" *Paper presented at the annual meeting of the The Midwest Political Science Association, Palmer House Hilton, Chicago, Illinois*, April 20, 2006.
- Grubestic, T. (1997). *Low cost geo-demographics: Generation and interpretation of lifestyle cluster groups*. Cincinnati: ACSM; ASPRS; AM/FM International; AAG; URISA; APWA.
- Halvorson, M. (2008). *Visual Basic 2008 step by step*. Redmond: Microsoft Press.
- Harris, R. (2005). *Geo-demographics, GIS and neighborhood targeting*. West Sussex: John Wiley & Sons.
- Jardine, D. (2003). Using GIS in alumni giving and institutional advancement. *Using Geographic Information Systems in Institutional Research*, 120, 77-89.
- MAP Light.org . (2008). *Remote Control*. Berkeley.
- MacEachren A. M., Gahegan M., Pike W., Brewer I., Cai G., Lengerich E, and Hardisty F 2004 Geovisualization for knowledge construction and decision-support. *Computer Graphics and Applications*, 24, 7-13.

- Mitchell, A. (2005). *The ESRI guide to GIS analysis, Volume 2: Spatial measurements and statistics*. Redlands: ESRI Press.
- Payton, R. and Moody, M. (2008). *Understanding philanthropy: Its meaning and mission*. Monroe: Indiana University Press.
- Rogerson, P. (2006). *Statistical methods for geography* (2nd ed.). London: Sage Publications.
- Unknown Author. (1999). *History and traditions*. Retrieved May 1, 2009, from The University of Redlands Web site: <http://www.redlands.edu/ux502.asp>
- Vest, C. (2006). *Industry, philanthropy, and universities: The roles and influences of the private sector in higher education* (Paper CSHE-12-06). Retrieved February 1, 2009, from Center for Studies in Higher Education: <http://repositories.cdlib.org/cshe/CSHE-12-06>

Appendix A. Mapping Statistics by States

Table A. Alumni Statistics Categorized by State

State	Alumni	Lifetime Giving (\$)	Avg. Lifetime Giving (\$)	Percentage that Donate (%)
AK	61	27,159.62	445.24	43
AL	38	8,308.00	218.63	39
AR	48	15,861.07	330.44	50
AZ	714	389,891.77	546.07	39
CA	24070	17,348,335.74	720.75	33
CO	544	407,033.99	748.22	41
CT	79	51,504.95	651.96	43
DC	49	83,188.65	1,697.73	53
DE	8	3,792.00	474.00	63
FL	242	102,564.84	423.82	31
GA	143	61,216.43	428.09	38
HI	326	203,366.22	623.82	53
IA	44	21,969.52	499.31	52
ID	162	5,1363.32	317.06	37
IL	206	166,555.20	808.52	43
IN	74	59,233.74	800.46	34
KS	61	28,283.07	463.66	31
KY	31	6,776.07	218.58	32
LA	24	14,826.25	617.76	33
MA	133	167,283.15	1,257.77	52
MD	141	141,362.41	1,002.57	45
ME	32	4,312.14	134.75	47
MI	90	5,6281.28	625.35	41
MN	130	5,7091.79	439.17	40
MO	105	21,665.57	206.34	40
MS	14	1,280.07	91.43	50
MT	75	72,129.01	961.72	41
NC	161	116,547.45	723.90	39
ND	7	4,009.93	572.85	71
NE	35	7,356.00	210.17	31
NH	23	10,402.20	452.27	65
NJ	91	12,3431.21	1,356.39	51
NM	150	273,161.99	1,821.08	47
NV	356	124,847.42	350.70	35
NY	211	180,369.19	854.83	44

OH	124	228,217.64	1840.46	43
OK	53	78,638.04	1,483.74	47
OR	636	48,5978.61	764.12	44
PA	138	105,315.82	763.16	45
RI	5	5,662.00	1,132.00	20
SC	51	35,072.57	687.70	41
SD	16	1,189.00	74.31	38
TN	97	15,411.69	158.88	37
TX	563	511,375.32	908.30	37
UT	161	50,504.35	313.69	34
VA	277	452,033.02	1,631.89	44
VT	24	33,96.07	141.50	33
WA	859	552,455.47	643.14	45
WI	68	39,473.39	580.49	50
WV	13	11,515.00	885.77	46
WY	35	12,045.49	344.16	46